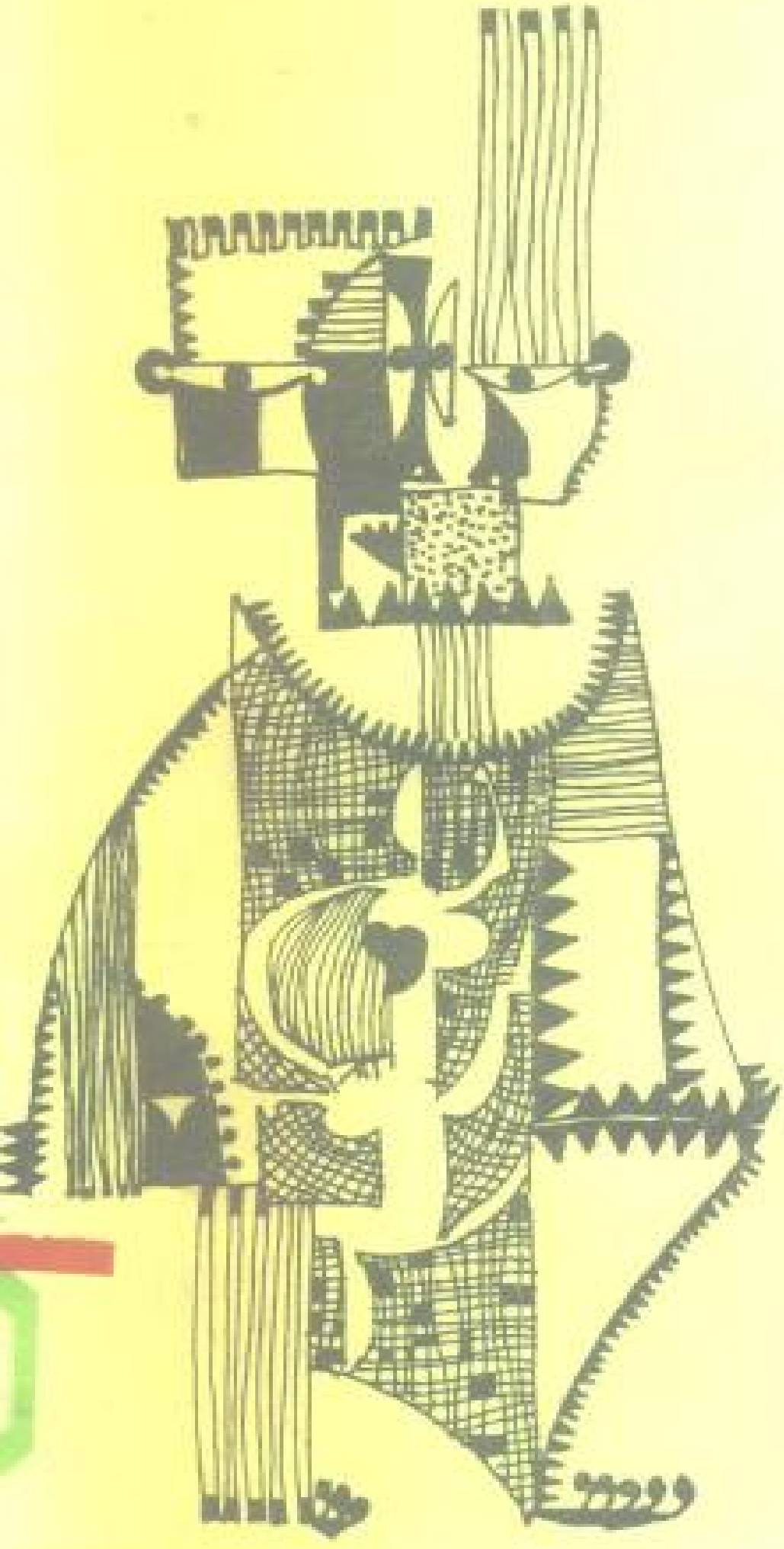


一条永恒的金带

VISIT…

2 022 3438 6

GEB—一条永恒的金带

四川人民出版社

一九八三年·成都

特约编辑：金观涛

责任编辑：罗由沛
封面设计：刘　溢
插　　图：刘　溢
　　曹　力
　　戴士和

走向未来丛书

GEB——一条永恒的金带　　乐秀成改写

四川人民出版社出版　（成都盐道街三号）
四川省新华书店发行　　四川新华印刷厂印刷

开本787×960毫米1/32　印张7　插页21　字数105千
1984年1月第一版　　1984年1月第一次印刷
印数：1—16,500册

书号：2118·13　　定价1.30元

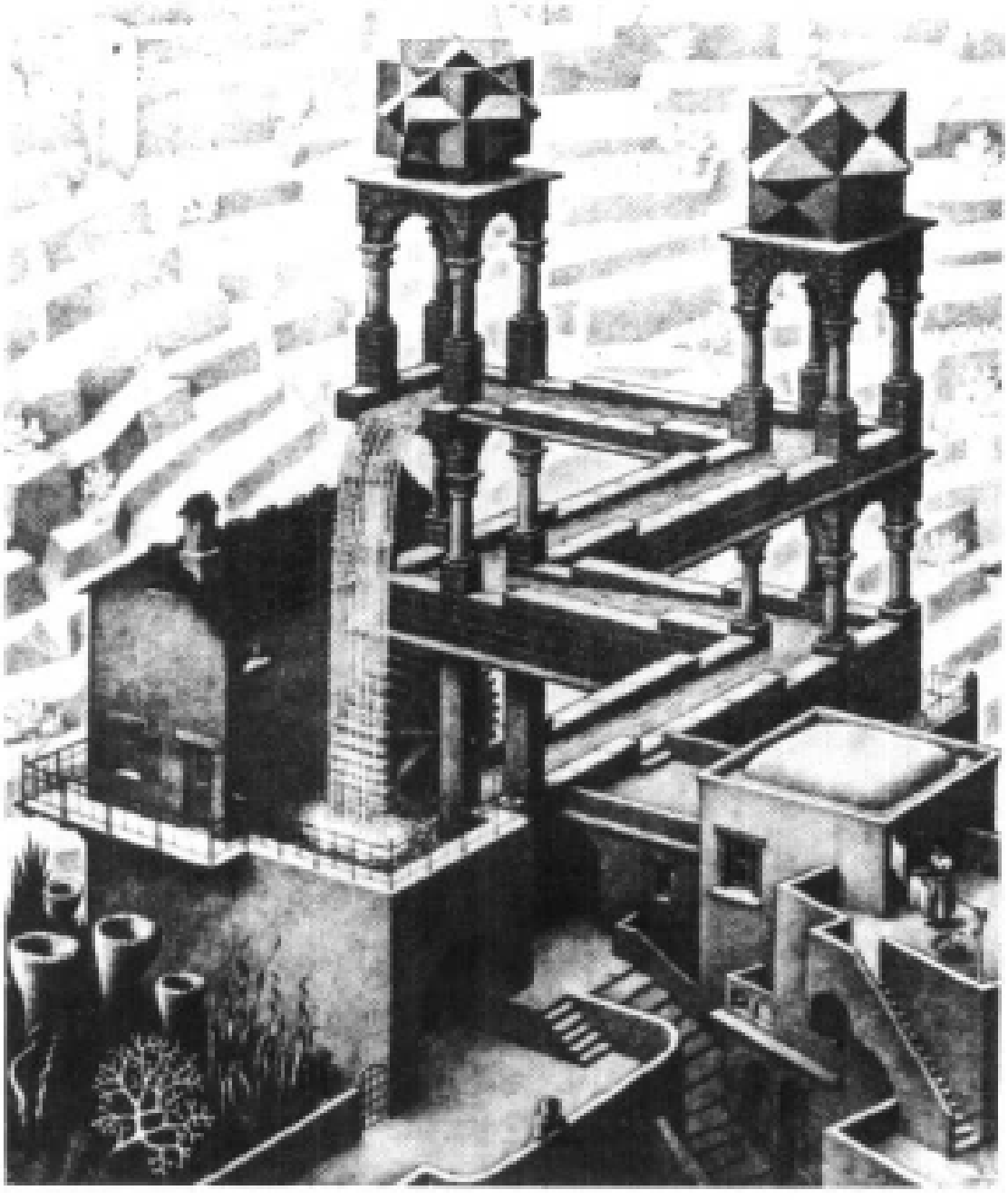


图1　瀑布　　M·C·埃舍尔（版画，1961）

瀑布一泻而下，汇集到池子中，然后顺着水渠往下流去。可是拐了几道弯之后，突然又回到了瀑布口！真是不可思议。可是在画面上却表现得明明白白，天衣无缝。这就是怪圈。

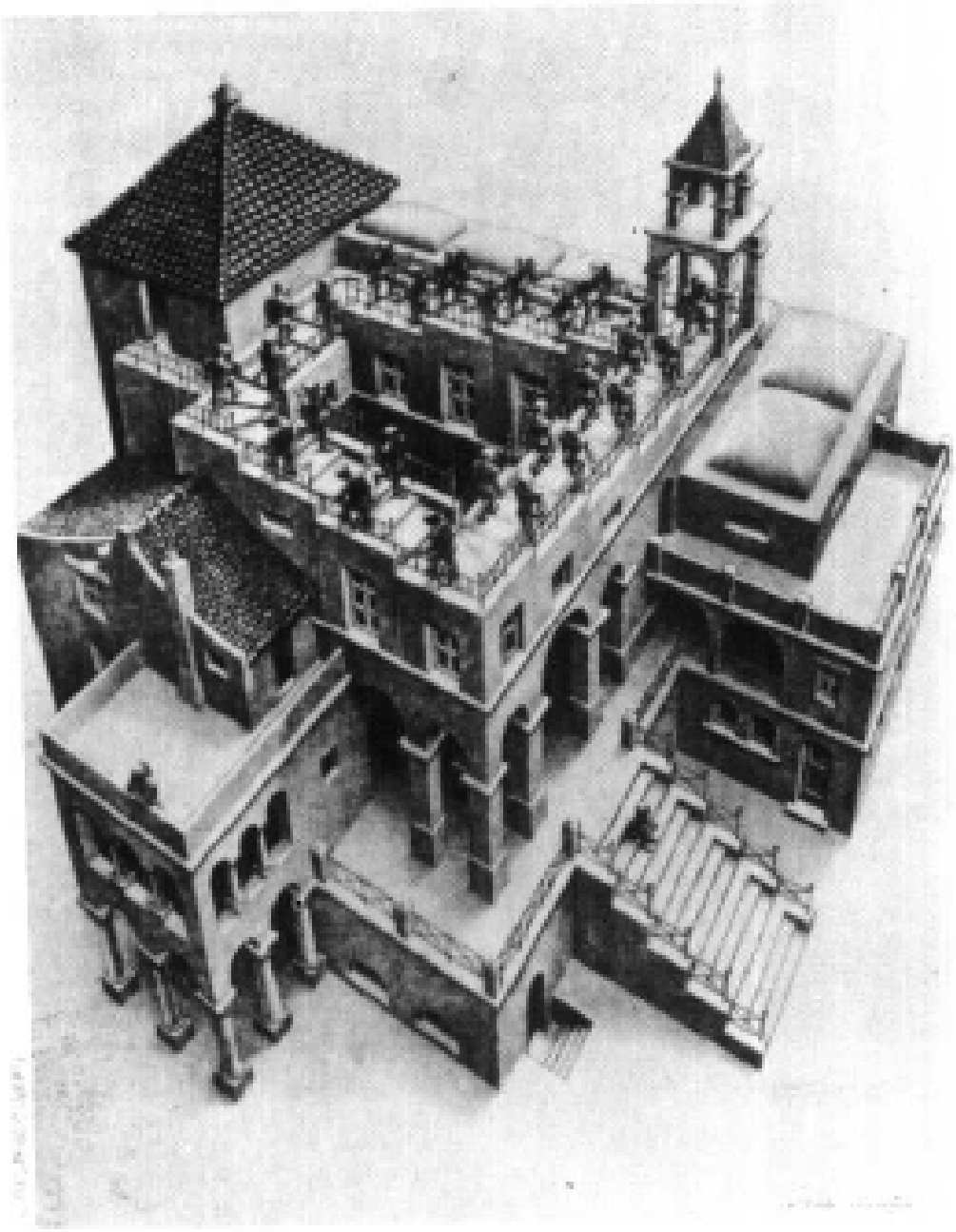


图2　上升与下降　　M·C·埃舍尔(版画，1960)

一队僧侣沿着楼梯向上走，拐了四次后又回到了原处。另一队僧侣沿着楼梯向下走，最后也退回到原处。而且两队人走同样的楼梯。这又是一个怪圈。

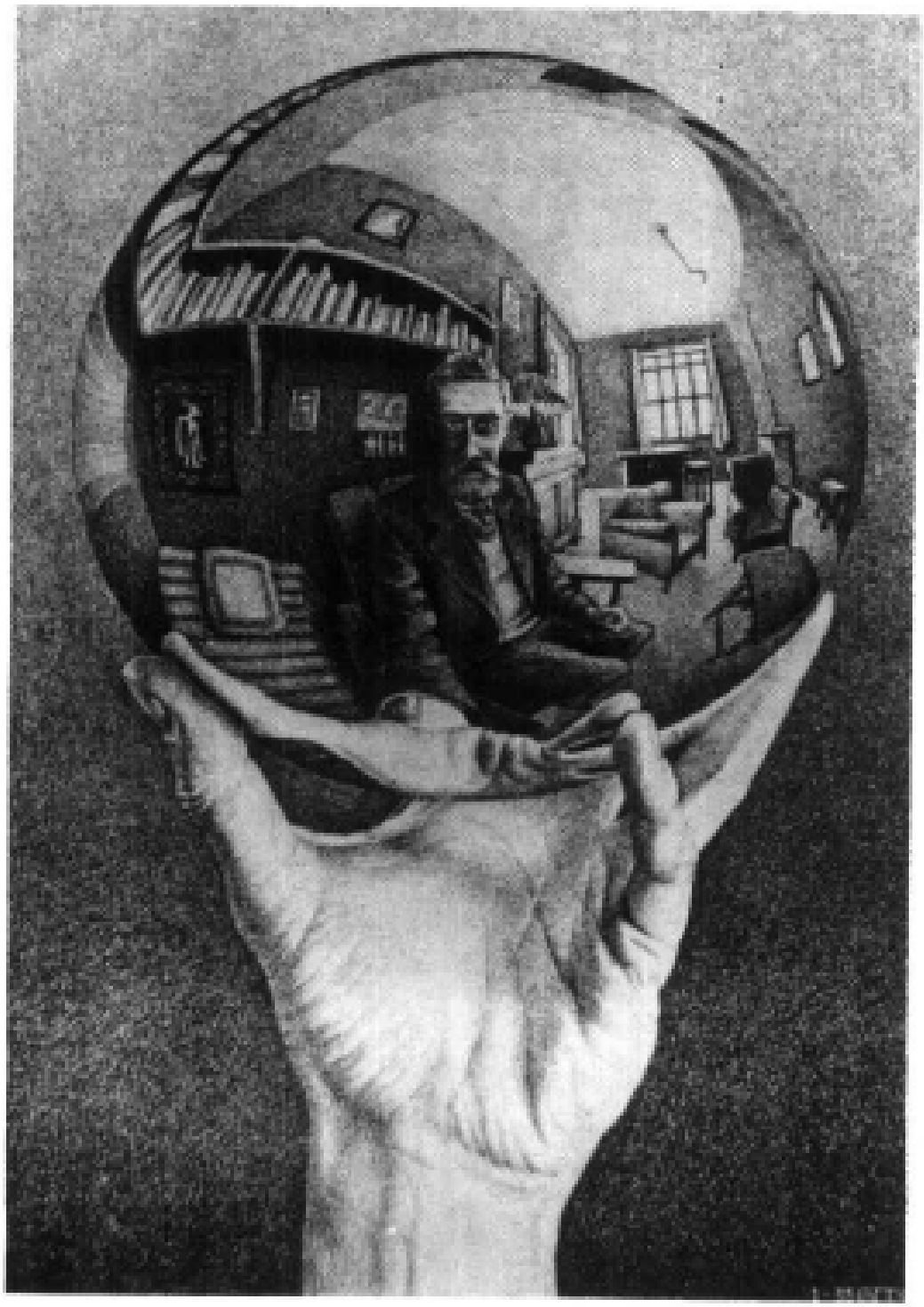


图3　举着反光球的手　　M·C·埃舍尔的自画象。（版画，1935）

球反射出举着球的画家，画家的书房。不过最妙的是我们从反光中看到了球本身。这难道不也是一个怪圈吗？

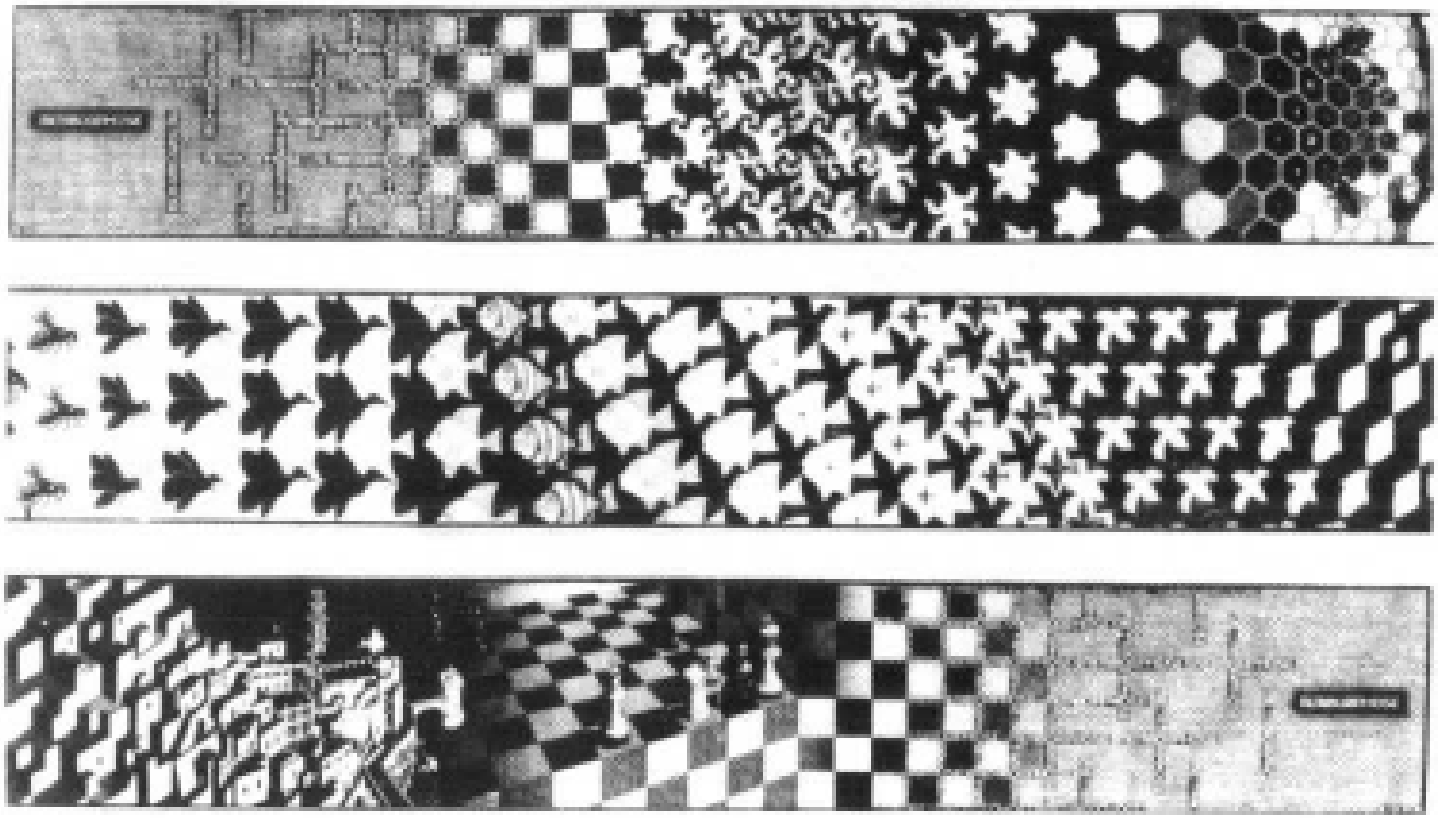


图4　变形　　M·C·埃舍尔(木刻,19.5cm×400cm, 1939-1940)

这是一幅围成圈、首尾相接的长画。你可以随着画面的逐渐变化慢慢往前走，不知不觉地回到了原处。你可以永远这样走下去。

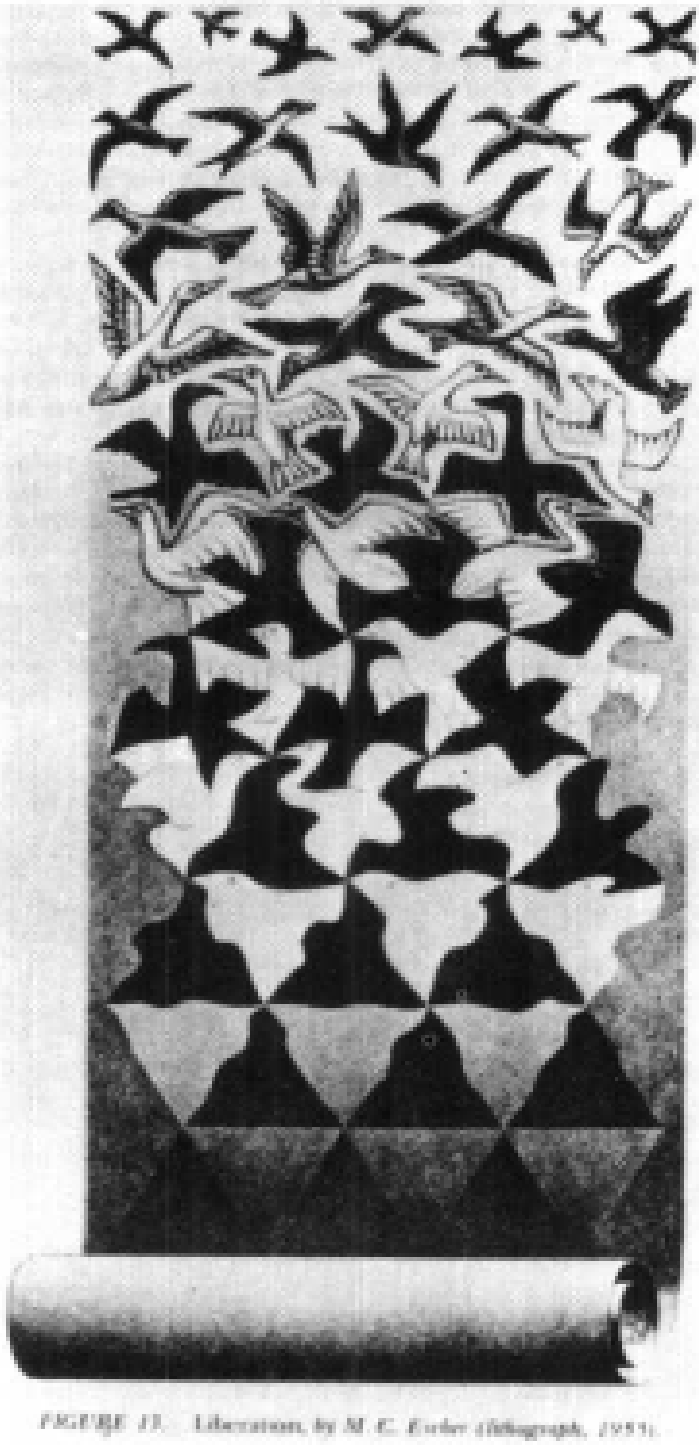


FIGURE 11. Liberation, by M. C. Escher (lithograph, 1955).

图5　解放　　M·C·埃舍尔（版画，1955）

上方是自由翱翔的鸟，下面是死板规则的几何图形，中间则是花哨的过渡区域。形式力图把握活生生的现实。现实却象鸟儿一样要求智能的解放。

图6　镶嵌图案　　M·C·埃舍尔（版画，1957）

白色的图形与黑色的图形镶嵌在一起，它们构成亲密无间的整体。它们互补互衬，互为背景。有谁能够想到这与数学定理也有密切的关系呢？

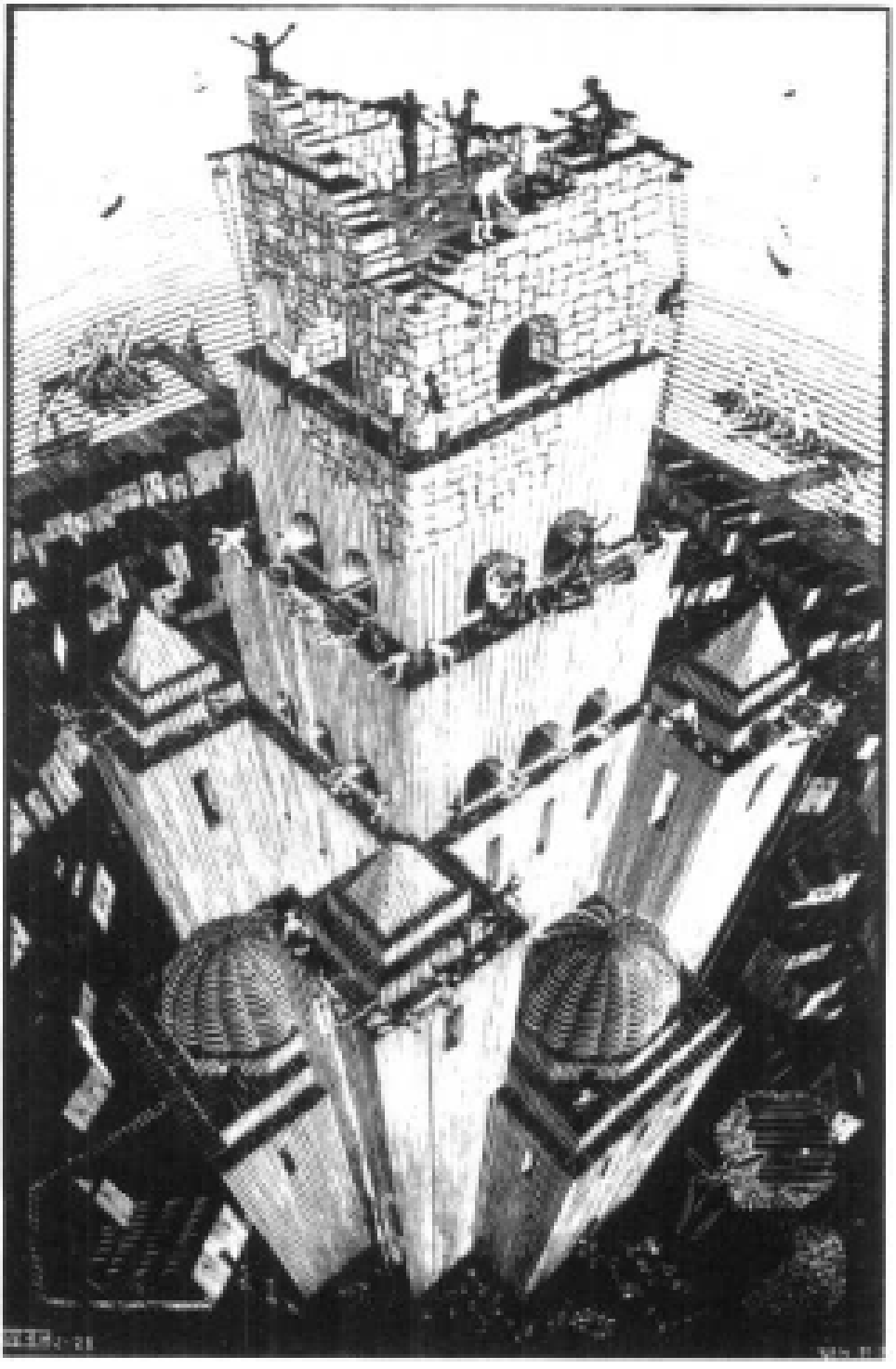


图7。 巴比仑塔 M·C·埃舍尔（木刻，1928）

高耸入云的巴比仑塔使上帝都感到敬畏。它象征着庞大、严谨的数学理论体系。它最后没有完全建成又给人以什么探测的启示呢？

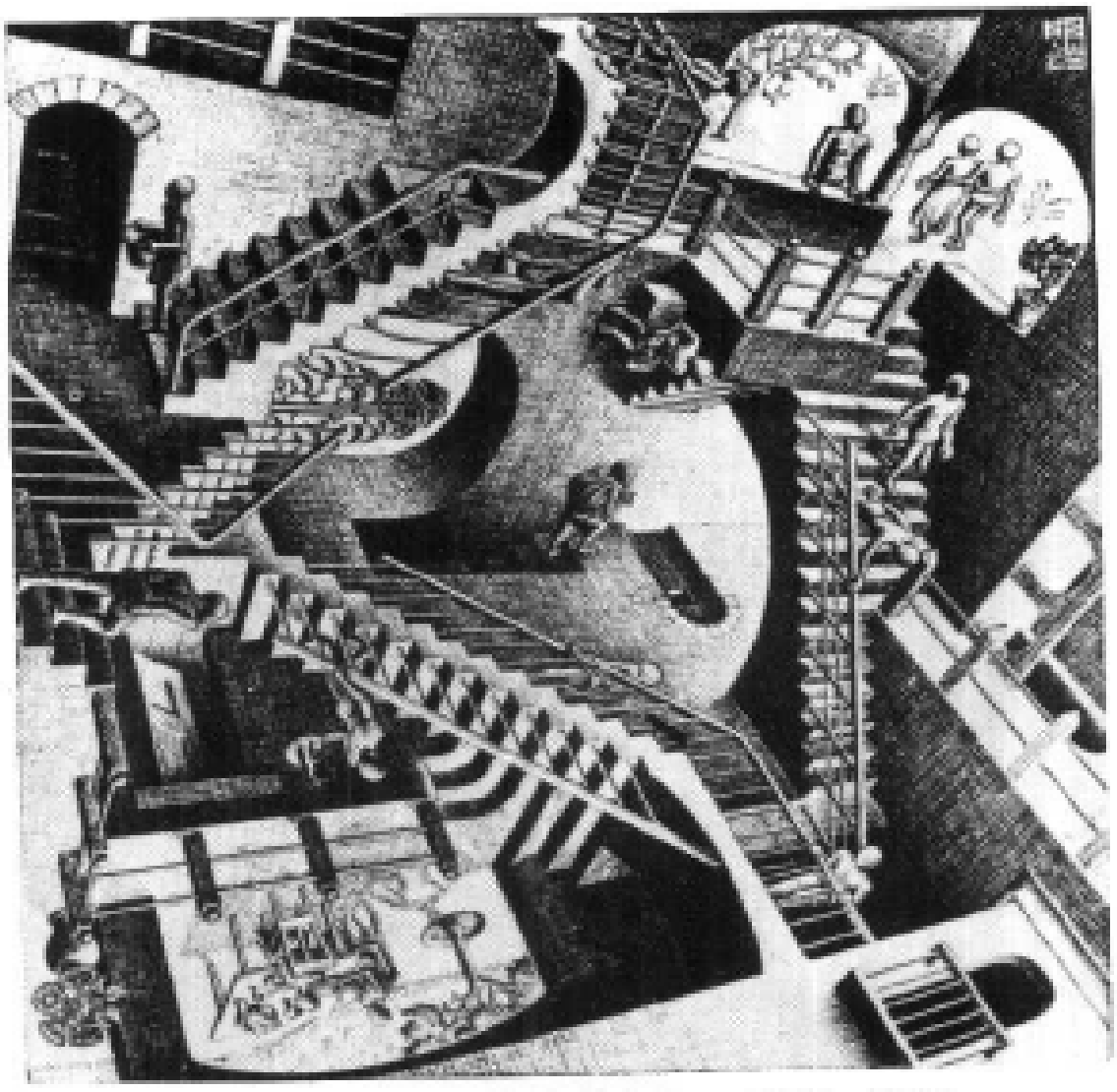


图8　相对性　　M·C·埃舍尔　《版画，1953》

以楼梯作为基点，根据它和邻近物体的相互关系，我们可以逐步认识周围的世界。可是不用走多远就会遇到明显的矛盾。换一个基点还是这样。这里究竟有什么地方不对头了呢？

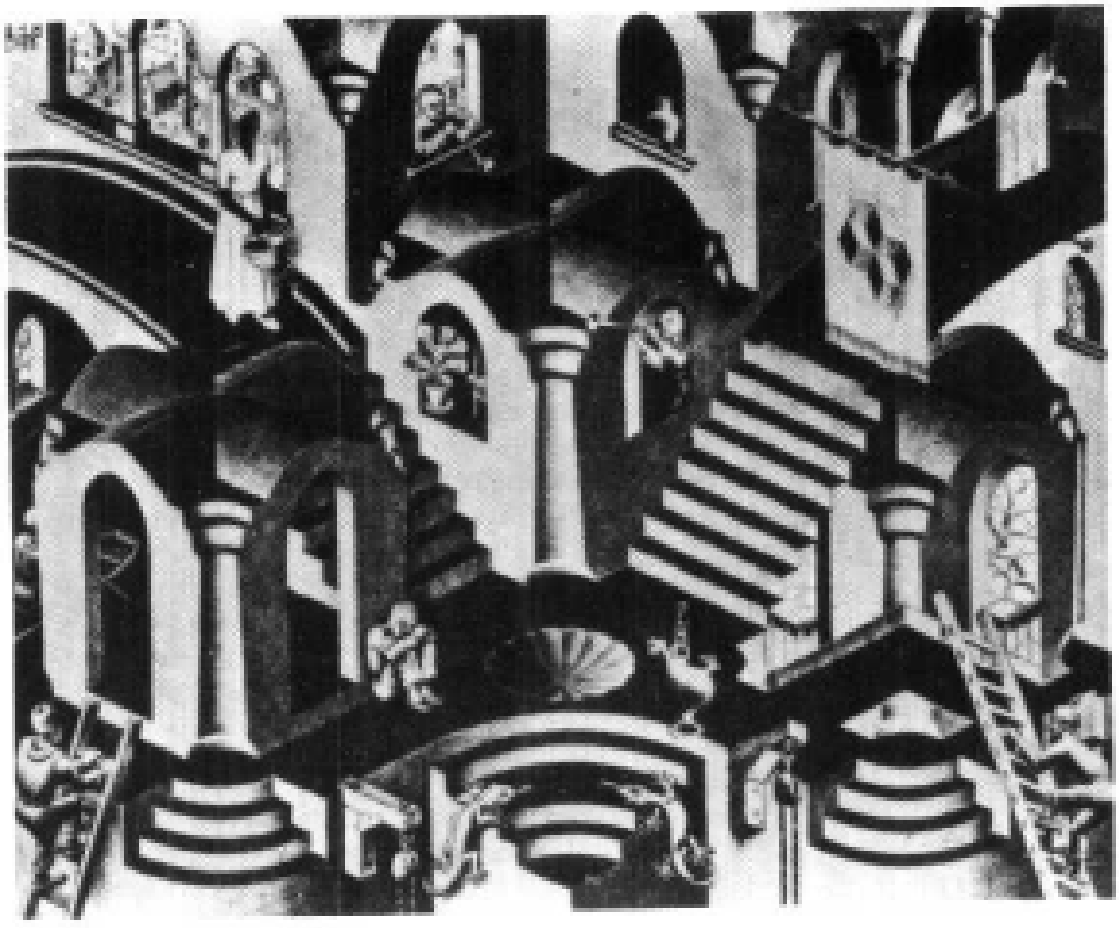


图9　凸与凹　　M·C·埃舍尔（版画，1955）

两个内部一致的世界共同构成了一个不一致的世界。难道你没有兴趣到这样的世界去周游一番？

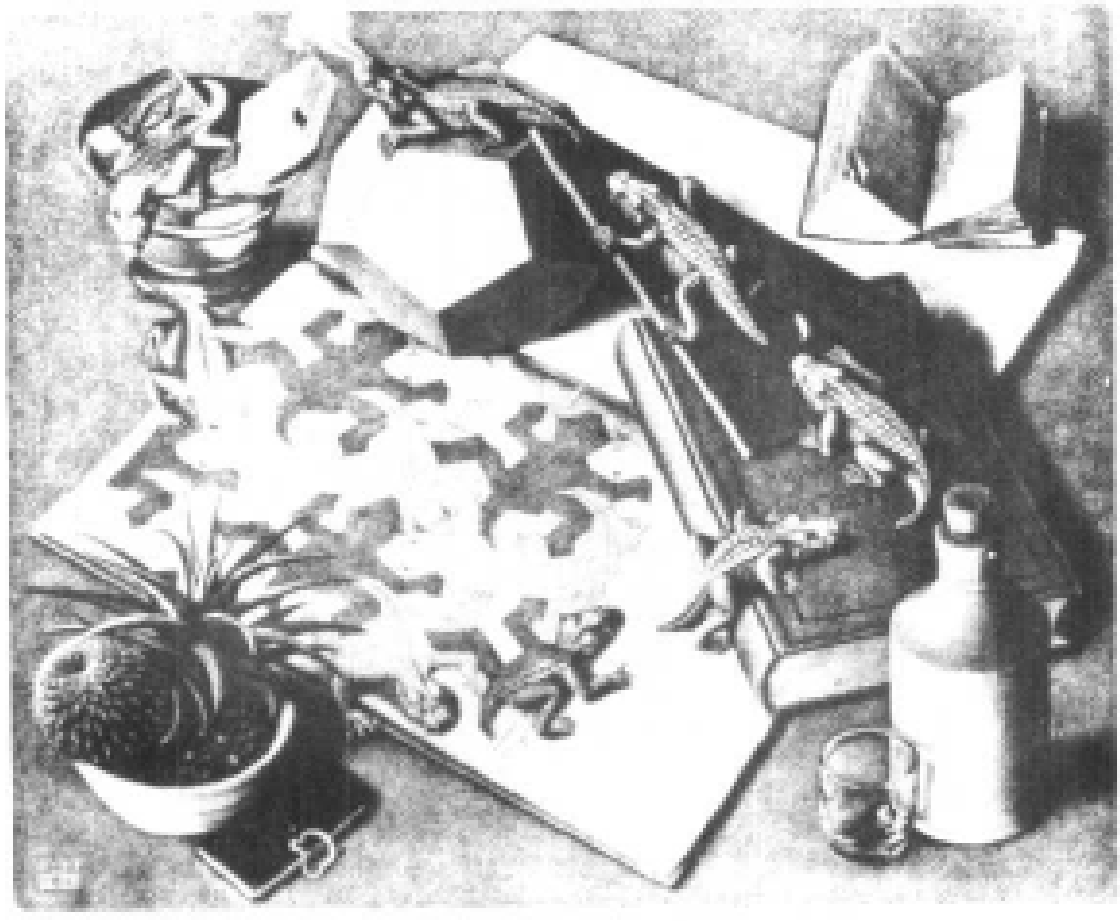


图10　蜥蜴　　M·C·埃舍尔（版画，1943）

同样是蜥蜴，一种是三维世界的动物，另一种是二维平面的图案。从一个世界进入另一个世界，也是从一个层次进入另一个层次。

图11　鱼和鳞　　M·C·埃舍尔　（木刻，1959）

每一片鳞是一条鱼，一种变形的相似，一种似乎不可能的相似。然而生物的遗传机制却为此提供了合理的解释。

图12　蝴蝶　　M·C·埃舍尔　（木雕，1950）

你能在这些蝴蝶中找到哪些共同点，从一只蝴蝶映入另一只蝴蝶保存了什么不变量？信息。重要的是保存信息的不变。

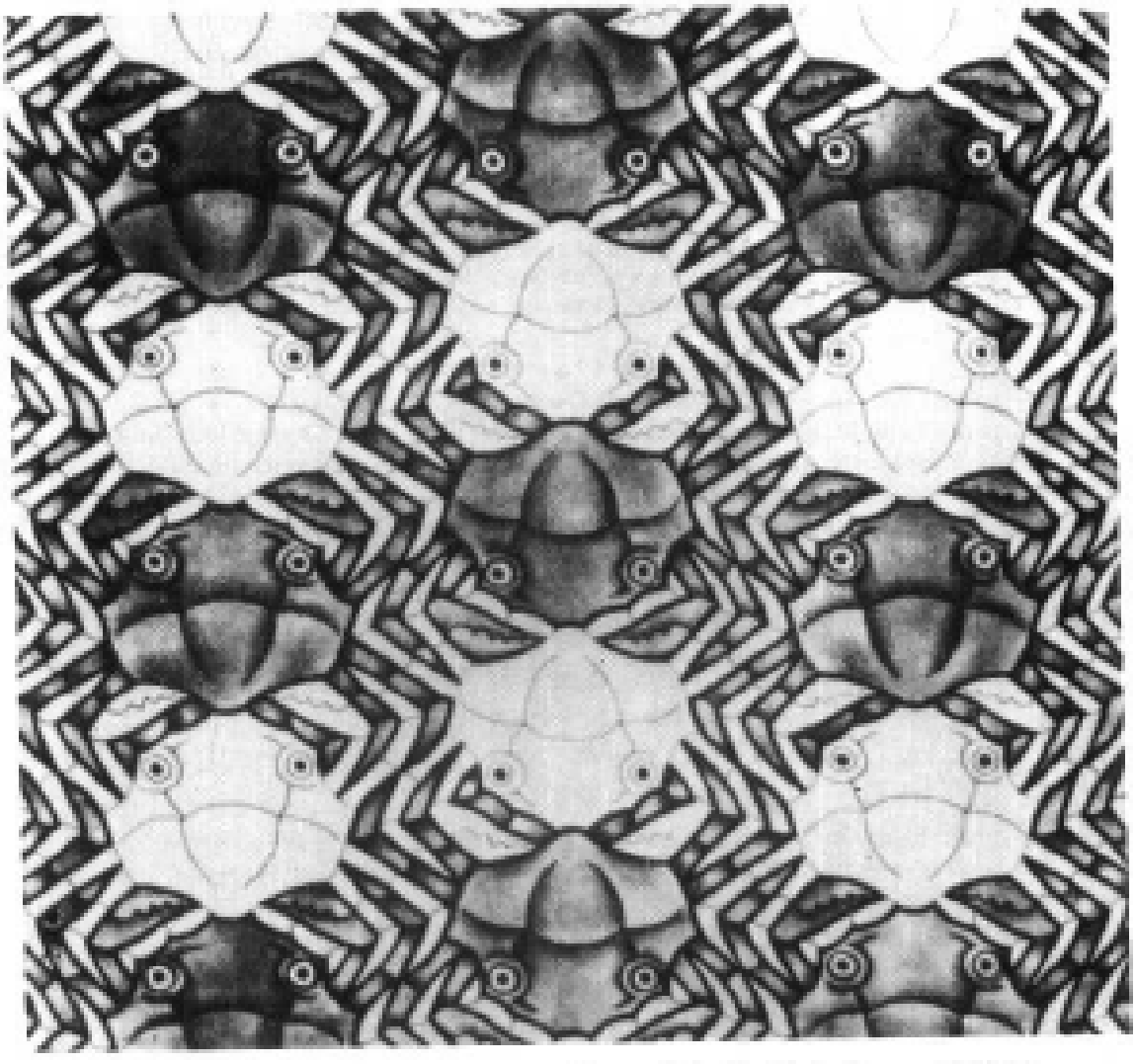


图13 “逆行(蟹式)卡农” M·C·埃舍尔 (1965)

这是一支看得见的乐曲，伴句的旋律与导句的旋律保持逆行的关系。然而一切是那样协调、自然。

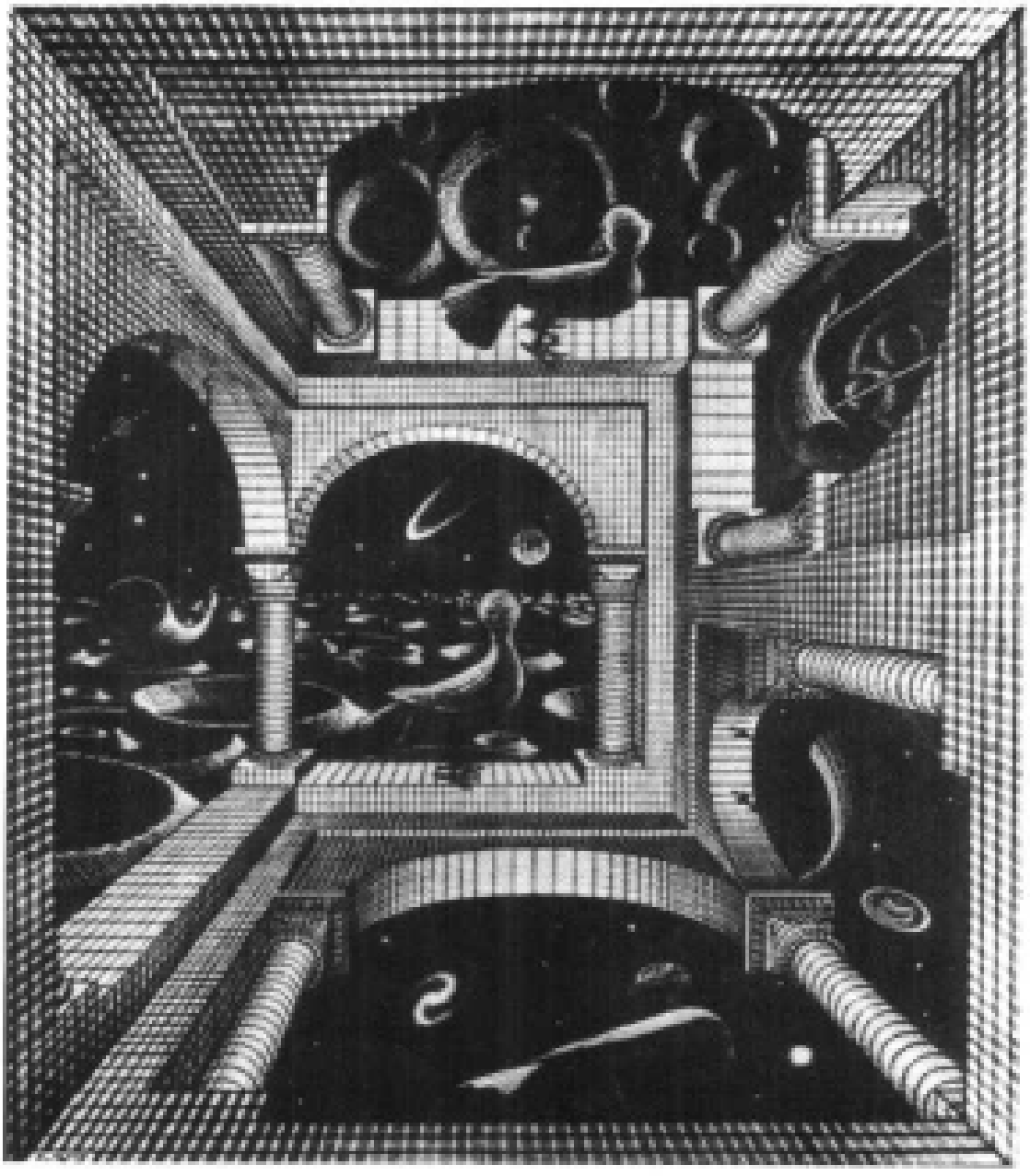


图14　另一个世界　　M·C·埃舍尔（木雕，1947）

用同样的线条表现真实和不真实的世界。

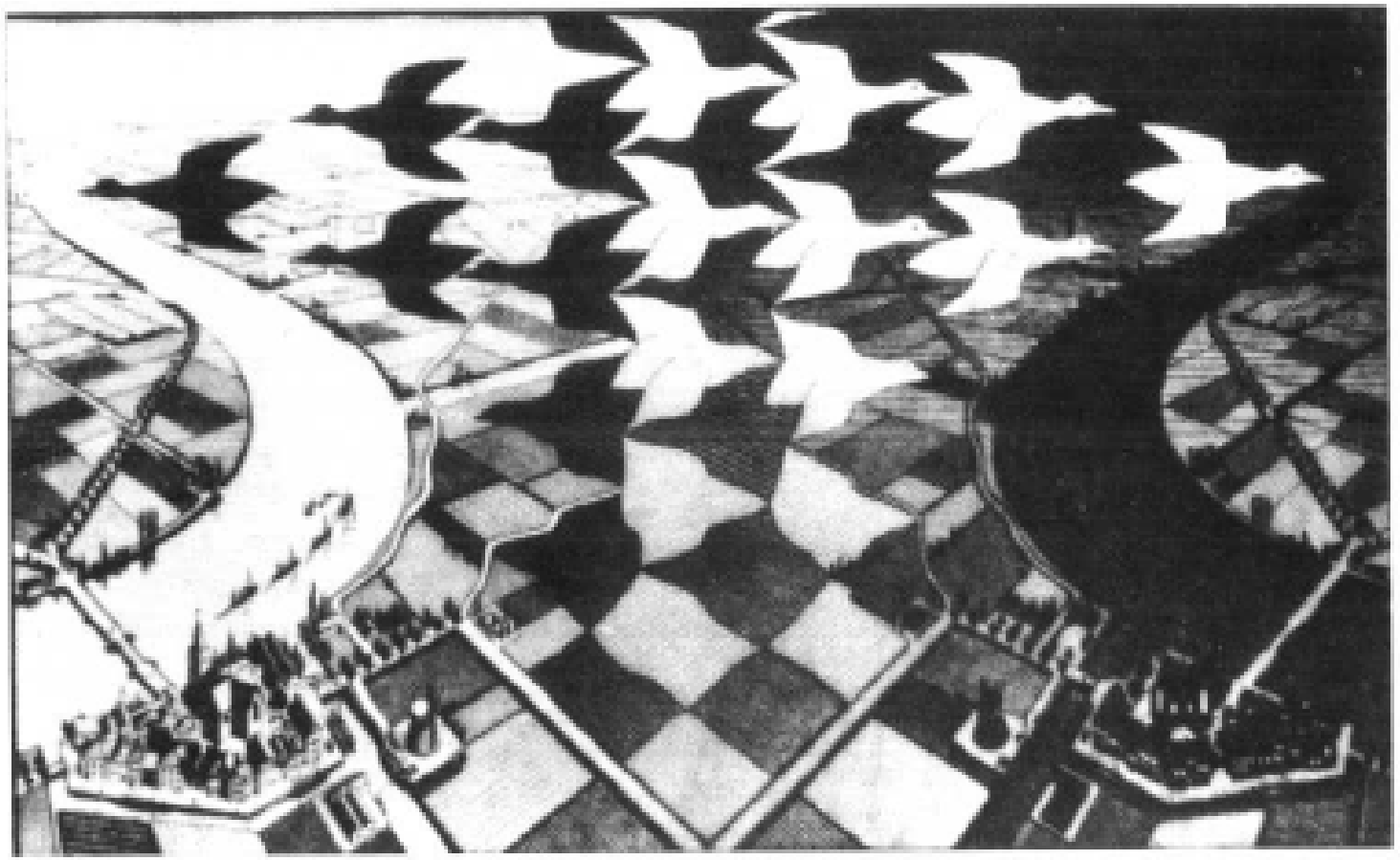


图15　白天和黑夜　　M·C·埃舍尔（木刻，1938）

这是真实的飞鸟和田野吗？这是真实的白天和黑夜吗？永恒的谜。

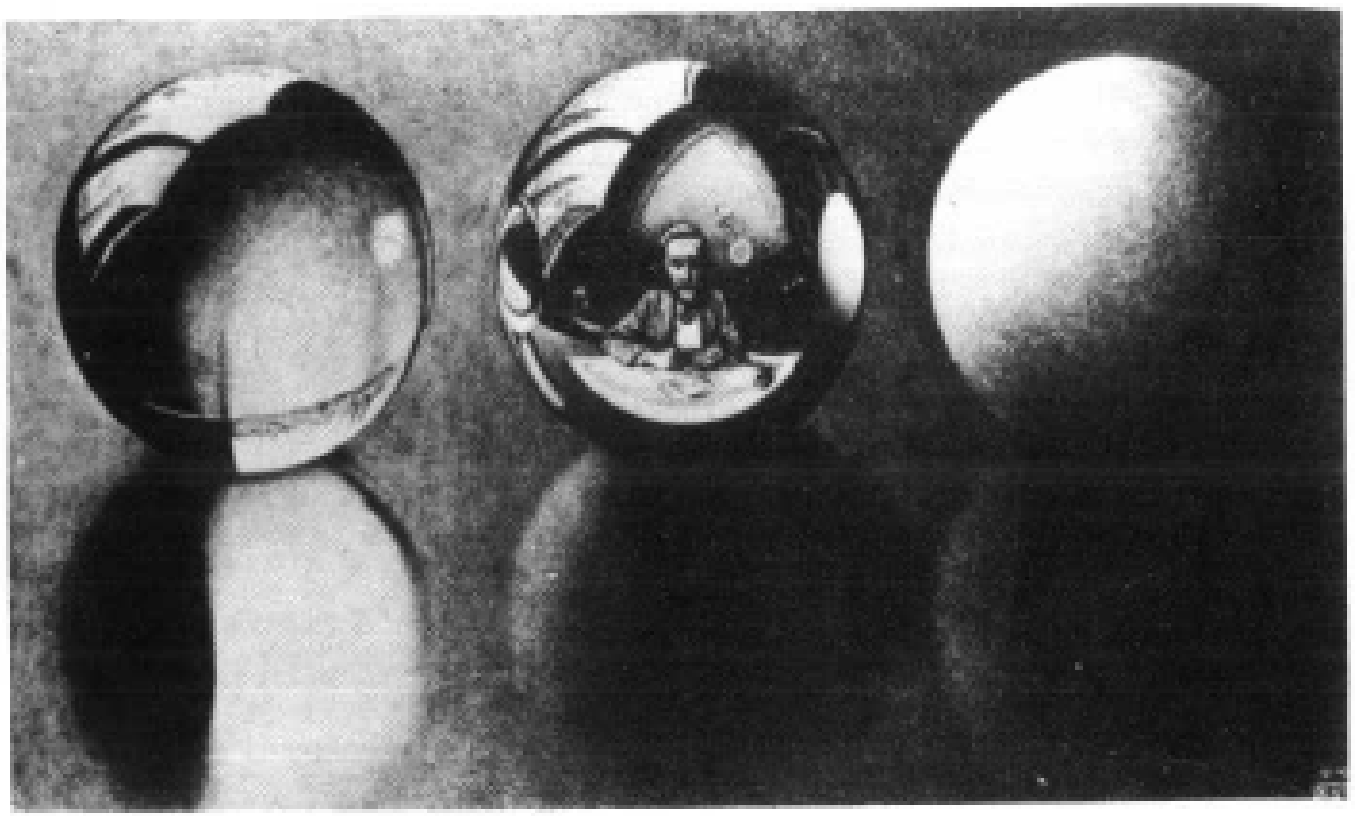


图16　三个球　　M·C·埃舍尔（版画，1946）

每个球以不同的方式反映周围的世界，它们也相互反映。世界就是一张相互紧密联结成一体的网。

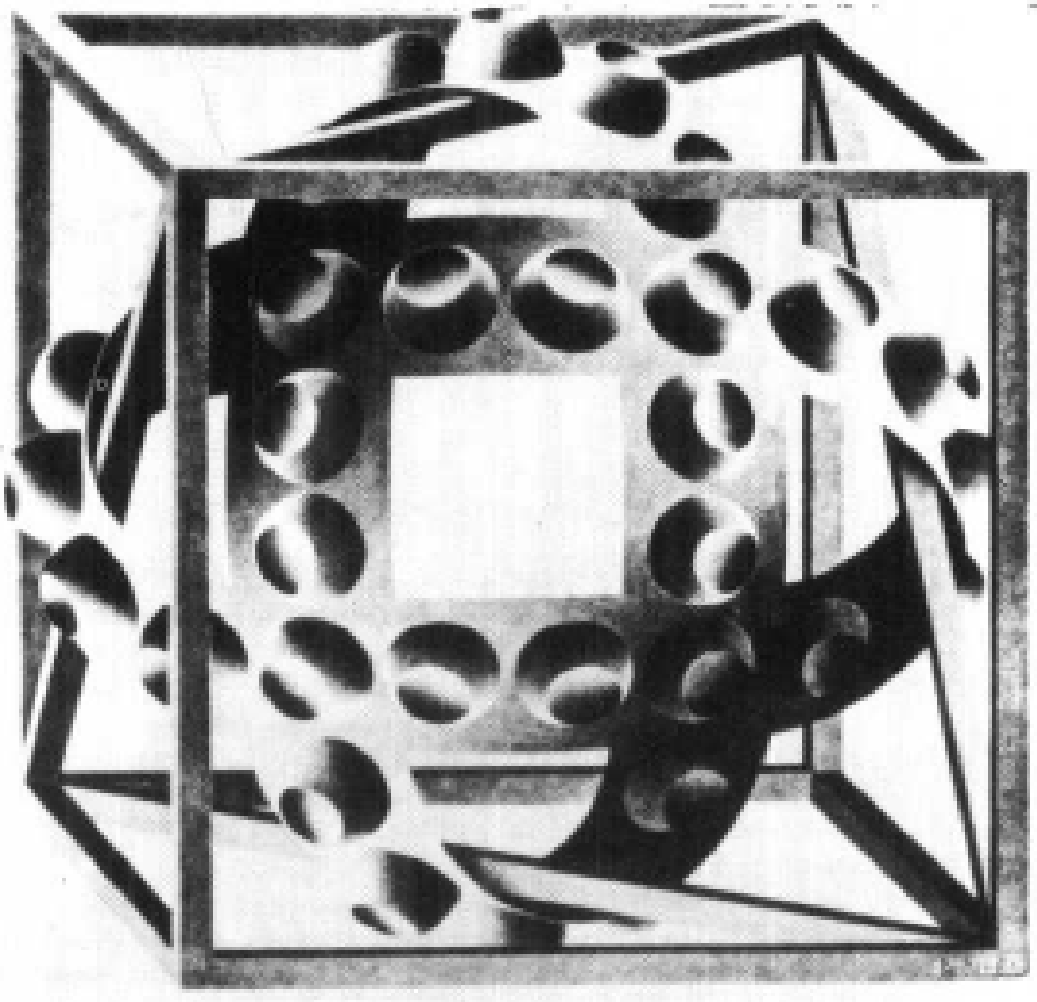


FIGURE 36. Cube with Magic Ribbons, by M. C. Escher (lithograph, 1957)

图17　带魔带的立方架　　M·C·埃舍尔（版画，1957）

魔带上的小泡这么看是放出来的，那么看是回进去的。玩弄这种魔术的又是谁呢？

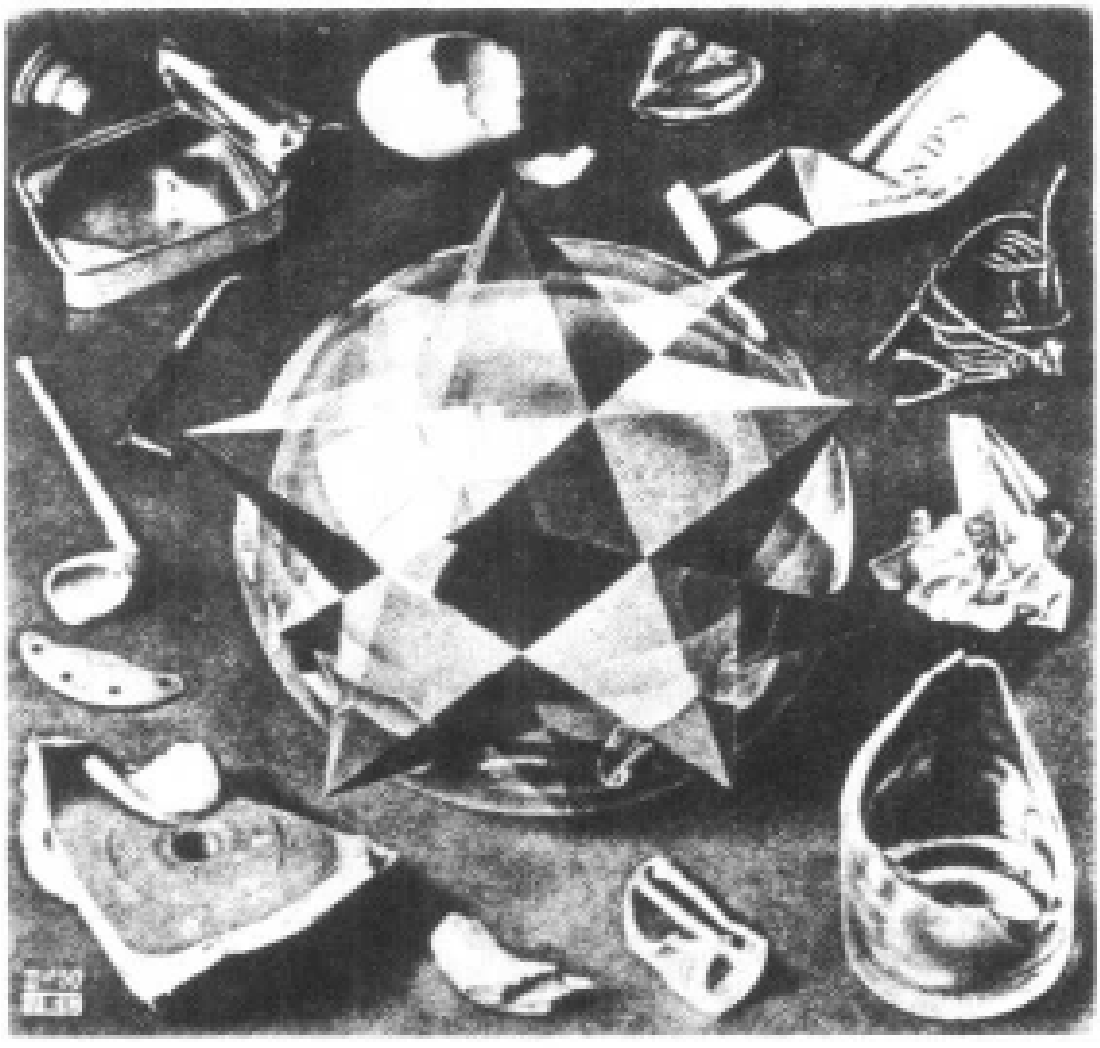


图18　有序和无序　　M·C·埃舍尔（版画，1959）

一串拉锯，却组成了别有风味的图案。无序中产生了有序。可以被认为最有序的数却总会表现一些无序的性质来。

图19　龙　　M·C·埃舍尔（木雕，1952）

一条咬住自己尾巴的龙。一条具有哥德尔精神的龙，一条拼命想从二维平面挣扎出来的龙。

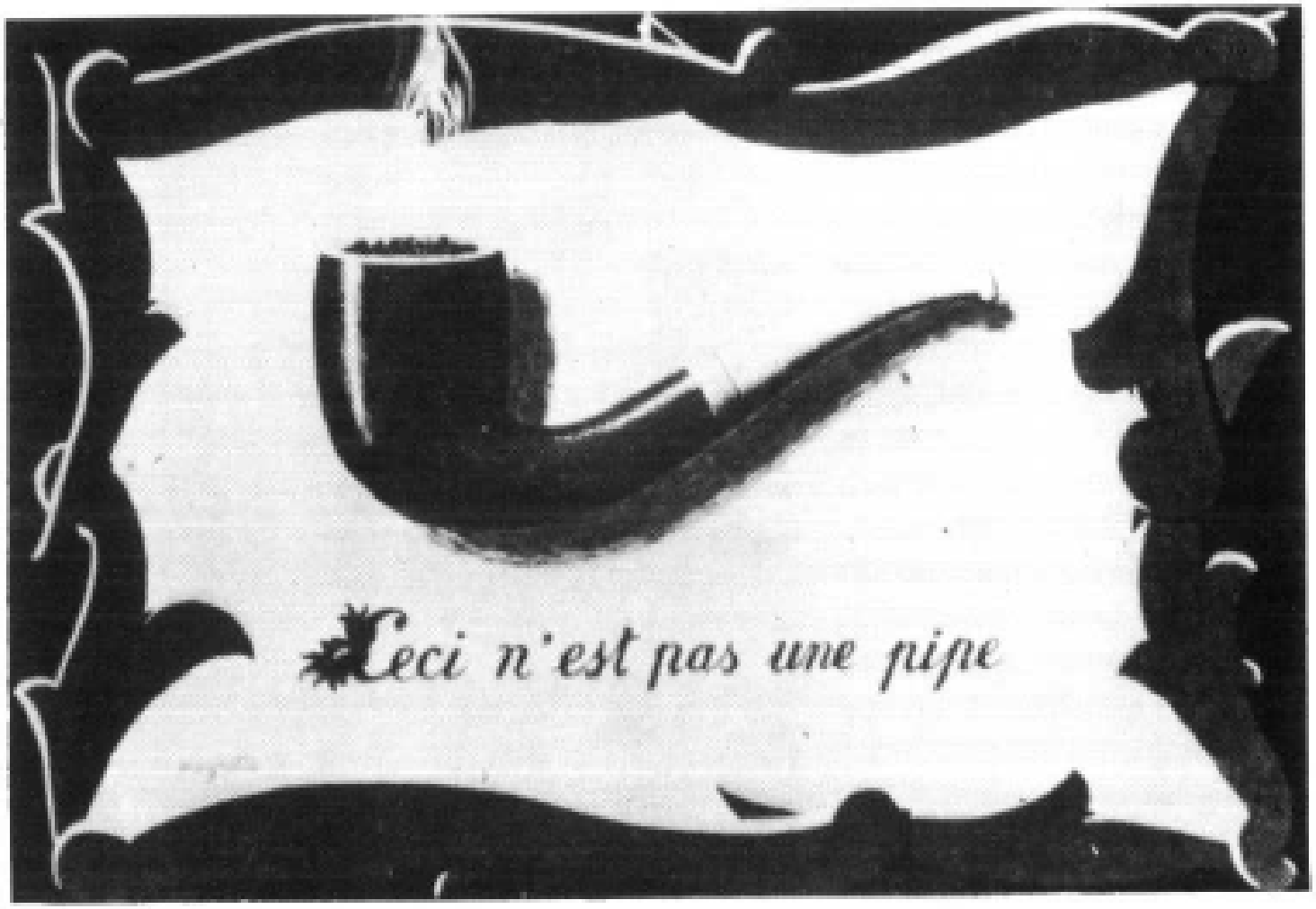


图20　曲调和歌　　勒内·马格里特（1964）

画中的一行法文：这不是烟斗。

“这”指什么呢？烟斗还是画？这幅画的语义解释会以哥德尔的方式自我互解。

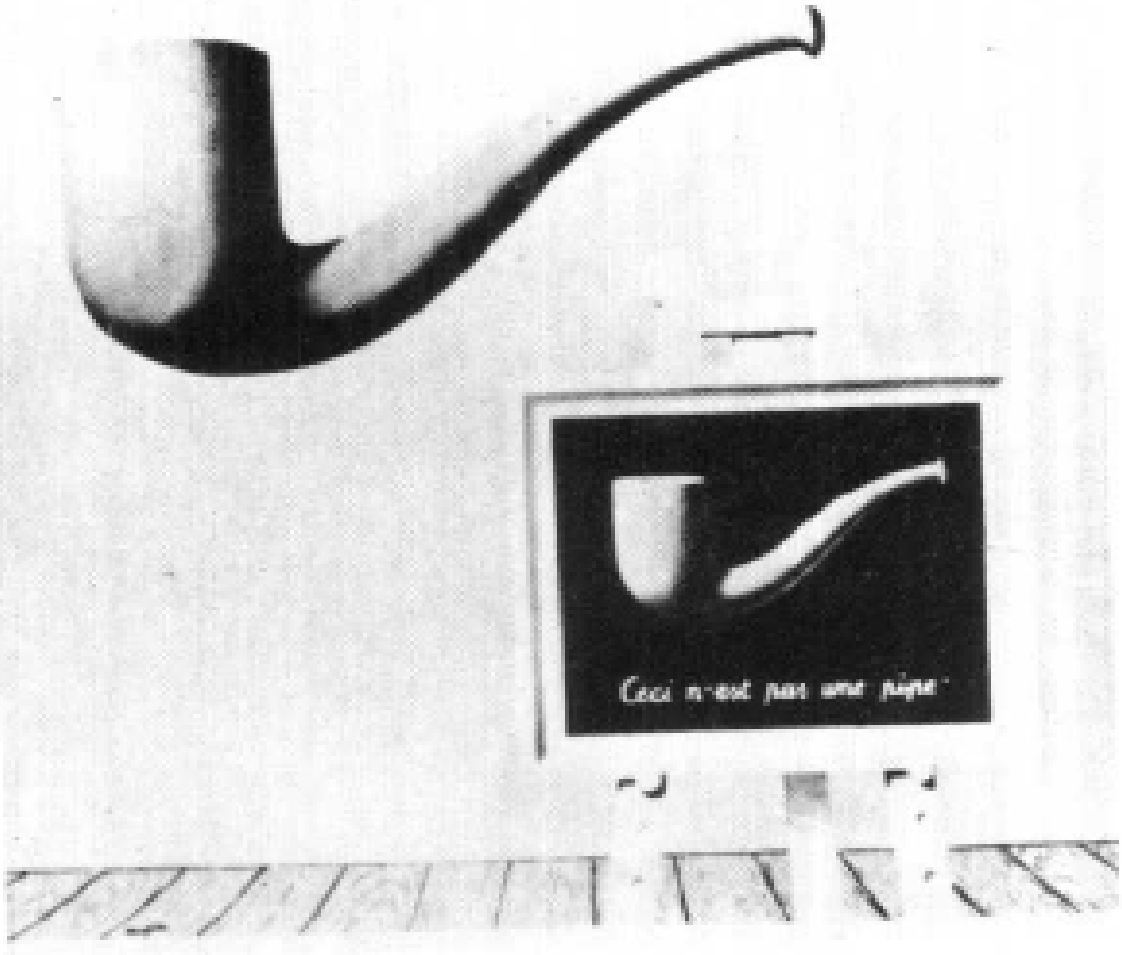


图21　两件神秘的东西　　勒内·马格里特　(1966)

画中的一行法文：这不是烟斗。

悬在空中的烟斗是否比画上的烟斗更加现实呢？又一幅语义画。

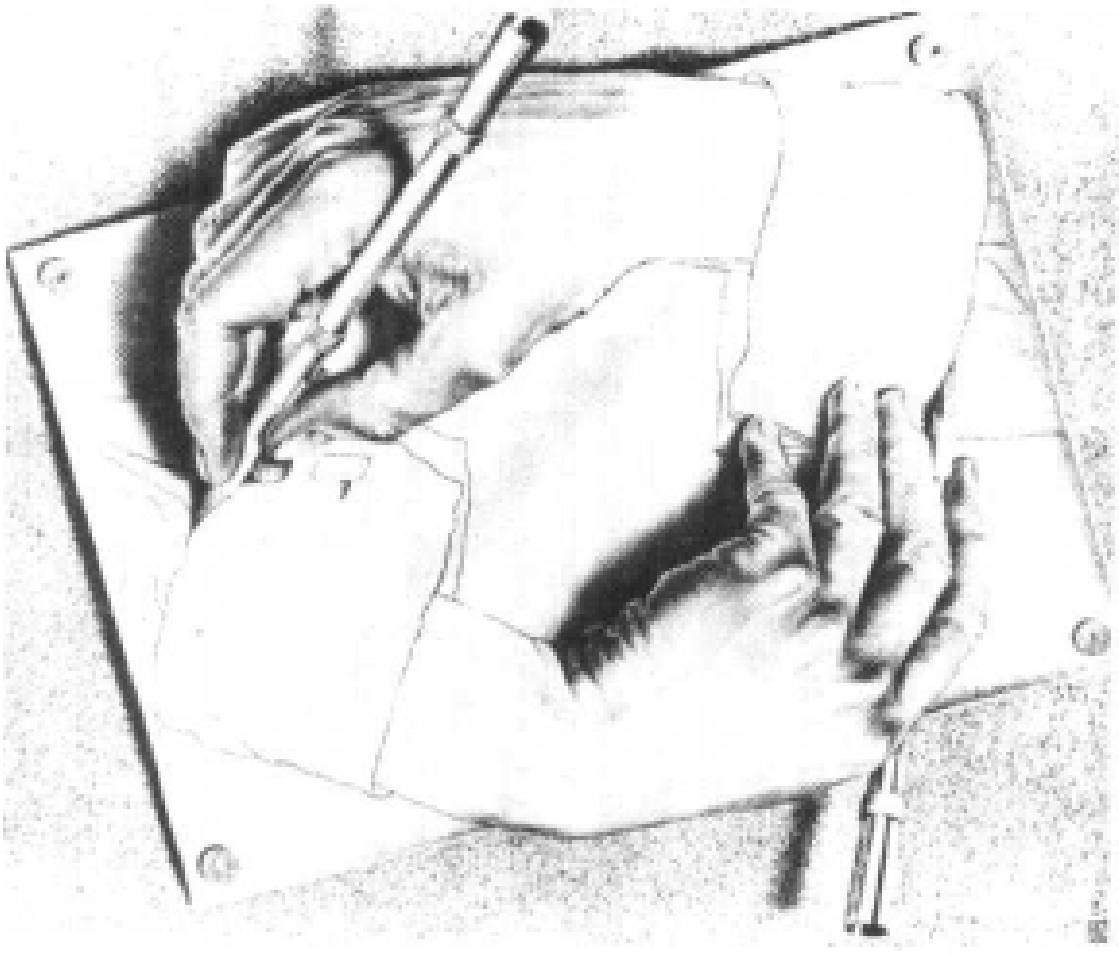


图22　画面的双手　　M·C·埃舍尔（版画，1948）

左手画右手，右手在画左手。它们到底又是谁画的呢？

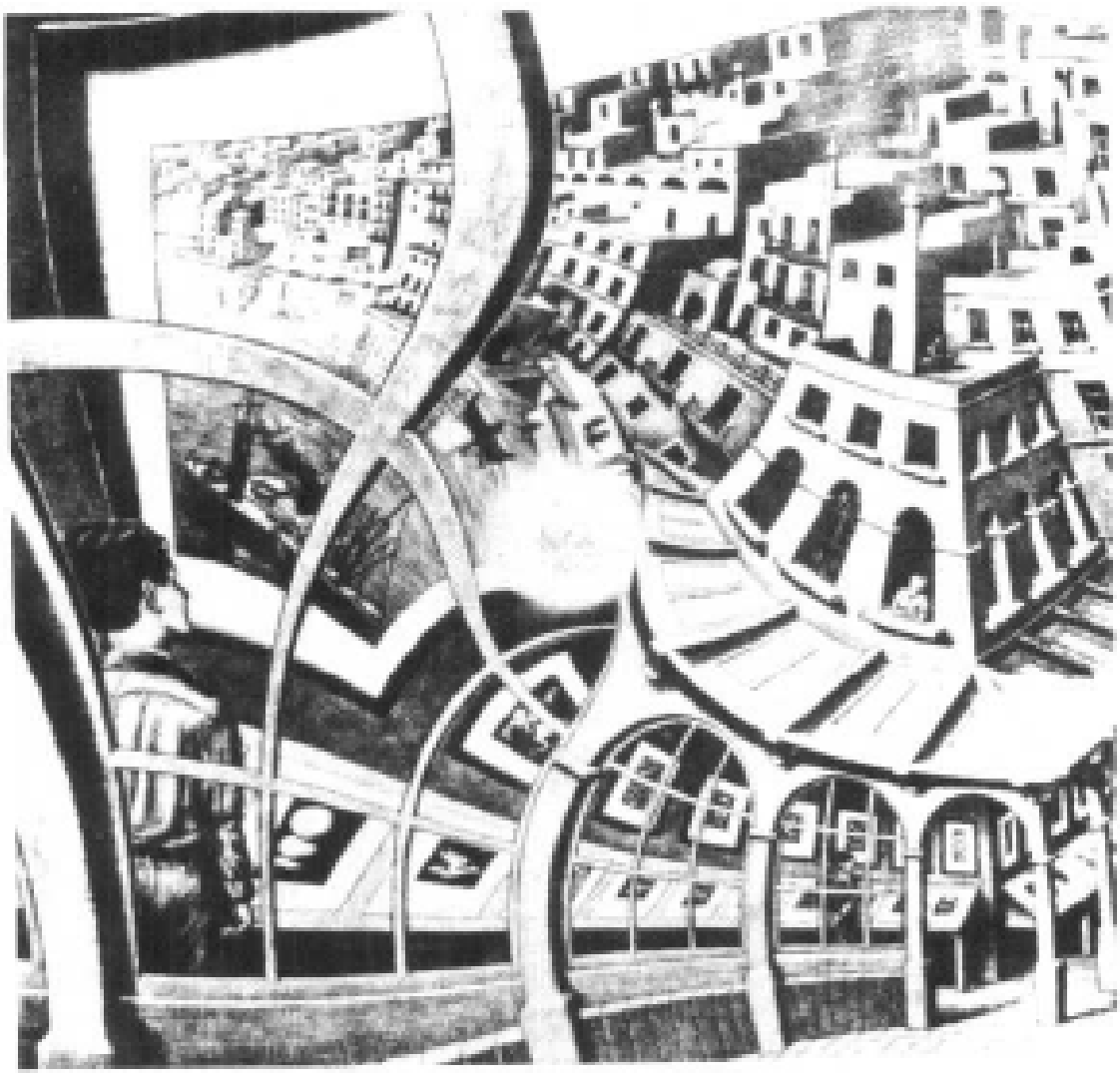
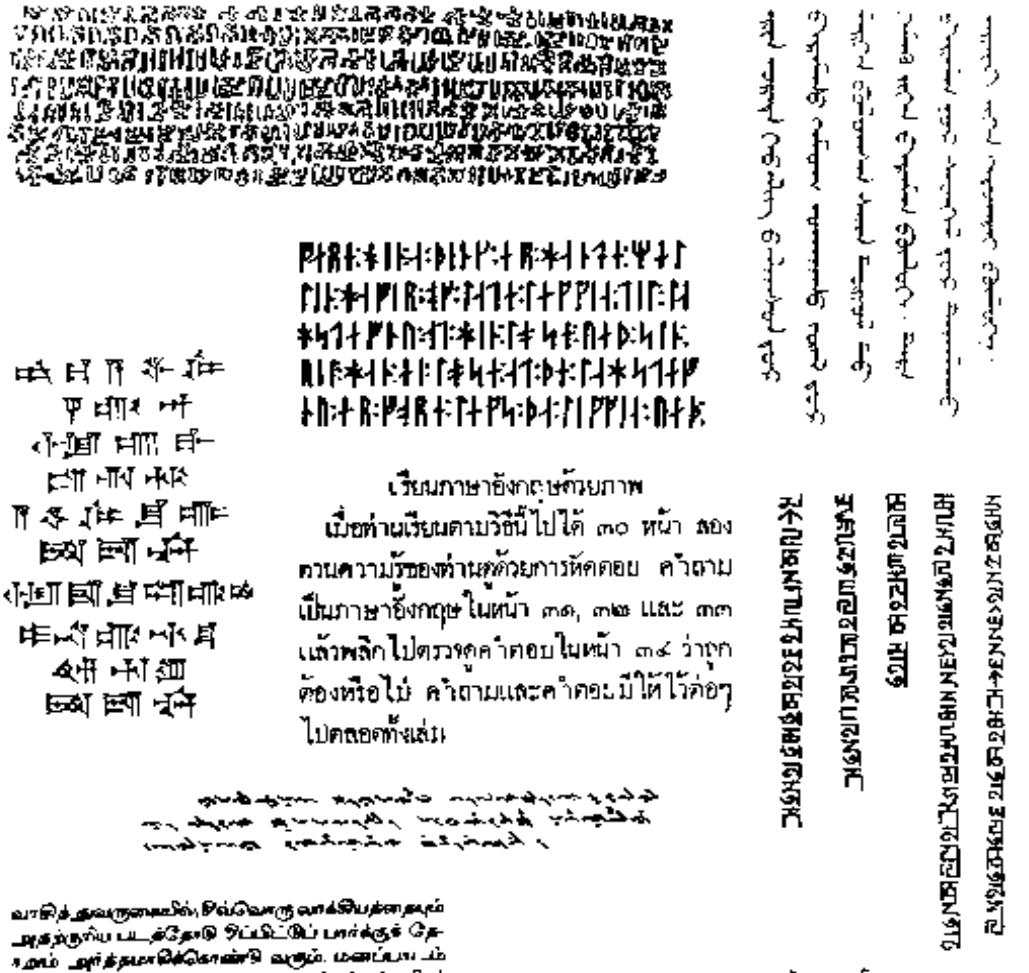


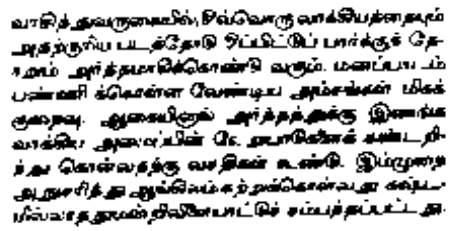
图23　画廊　　M·C·埃舍尔（版画，1956）

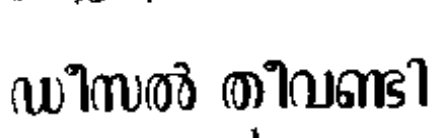
画中有画，城市中有城市，自我相关，最紧凑的怪圈。



เรียนภาษาอังกฤษด้วยภาพ

เมื่อท่านเรียนตามบทนี้ไปได้ ๓๐ หน้า ลองทดสอบความรู้ของท่านดูด้วยการหัดตอบ คำถามเป็นภาษาอังกฤษในหน้า ๓๑, ๓๒ และ ๓๓ แล้วพลิกไปตรวจดูคำตอบในหน้า ๓๔ ว่าถูกต้องหรือไม่ คำถามและคำตอบมีให้ไว้ต่อๆ ไปตลอดทั้งเล่ม





কত অজানারে জানাইলে তুমি,
কত ঘরে দিলে ঠাঁই
দূরকে করিলে নিকট বন্ধু,
পরকে করিলে ভাই।
পুরানো আবাস ছেড়ে যাই যবে
মনে ভেবে মরি কী জানি কী হবে,
নূতনের মাঝে তুমি পুরাতন
সে কথা যে ভুলে যাই
দূরকে করিলে নিকট বন্ধু,
পরকে করিলে ভাই

ഡീസൽ തിവണ്ടി എൻജിനു് നാടൻനിർമ്മിത വസ്തുക്കൾ

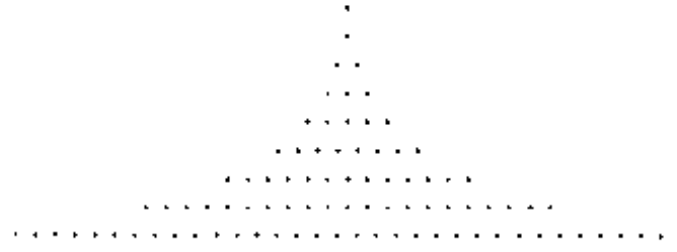
图24　字体集

formulated an anthropocentric concept. And of course, if contact were established with an alien civilization from another star system, we would feel supported in our belief that our own type of intelligence is not just a fluke, but an example of a basic form which reappears in nature in diverse contexts, like stars and uranium nuclei. This in turn would support the idea of meaning being an inherent property.

To conclude this topic, let us consider some new and old examples, and discuss the degree of inherent meaning which they have, by putting ourselves, to the extent that we can, in the shoes of an alien civilization which intercepts a weird object . . .

Two Plaques in Space

Consider a rectangular plaque made of an indestructible metallic alloy, on which are engraved two dots, one immediately above the another: the preceding colon shows a picture. Though the overall form of the object might suggest that it is an artifact, and therefore that it might conceal some message, two dots are simply not sufficient to convey anything. (Can you, before reading on, hypothesize what they are supposed to mean?) But suppose that we made a second plaque, containing more dots, as follows:



Now one of the most obvious things to do—so it might seem to a terrestrial intelligence at least—would be to count the dots in the successive rows. The sequence obtained is:

1, 1, 2, 3, 5, 8, 13, 21, 34,

Here there is evidence of a rule governing the progression from one line to the next. In fact, the recursive part of the definition of the Fibonacci numbers can be inferred, with some confidence, from this list. Suppose we think of the initial pair of values (1,1) as a "genotype" from which the "phenotype"—the full Fibonacci sequence—is pulled out by a recursive rule. By sending the genotype alone—namely the first version of the plaque—we fail to send the information which allows reconstitution of the phenotype. Thus, the genotype does not contain the full specification of

图25

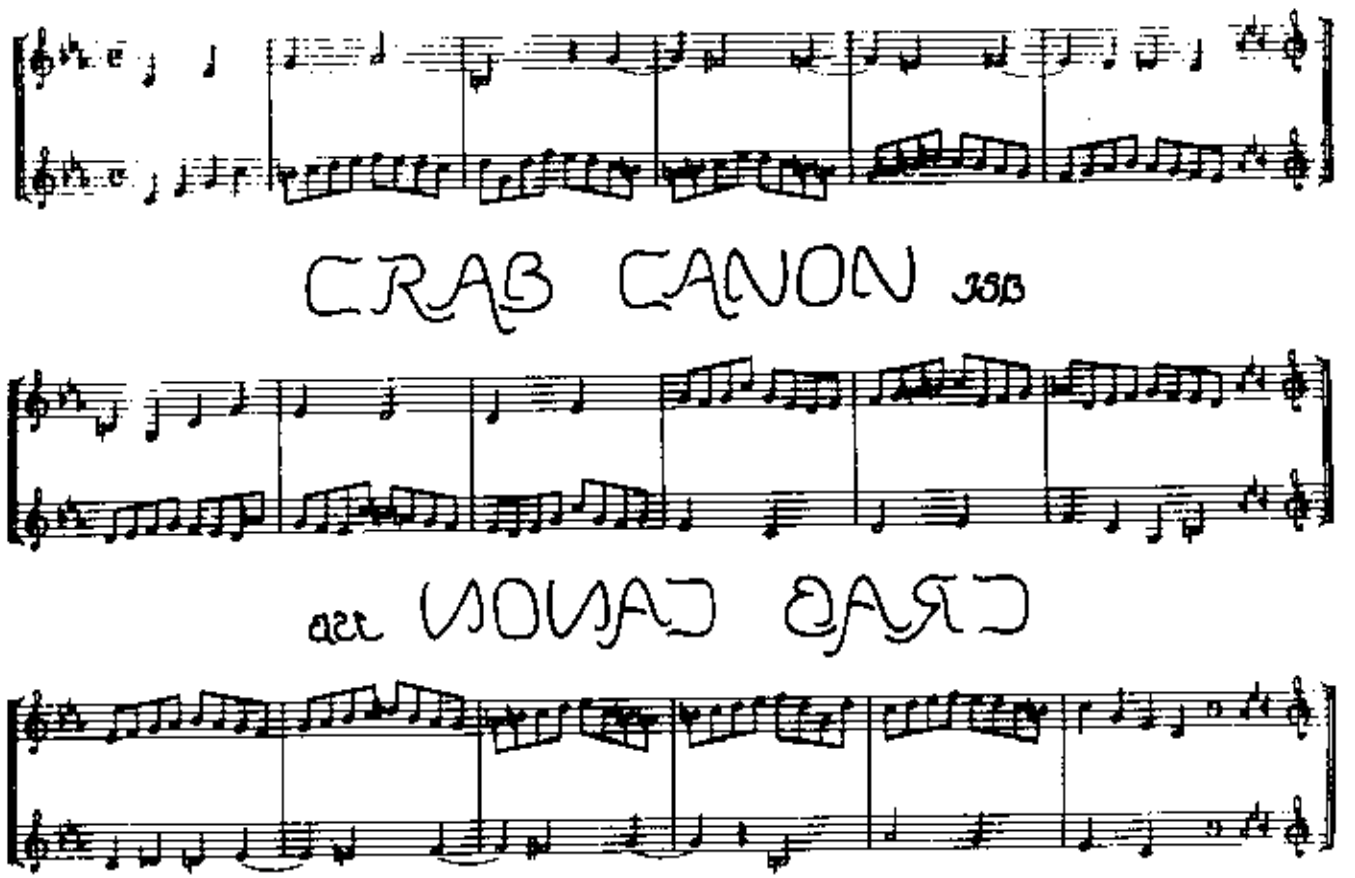



FIGURE 44. Crab Canon *from the* Musical Offering, *by J. S. Bach.* [*Music printed by Donald Byrd's program "SMUT."*]

图26　逆行卡农　　引自巴赫的《音乐的奉献》

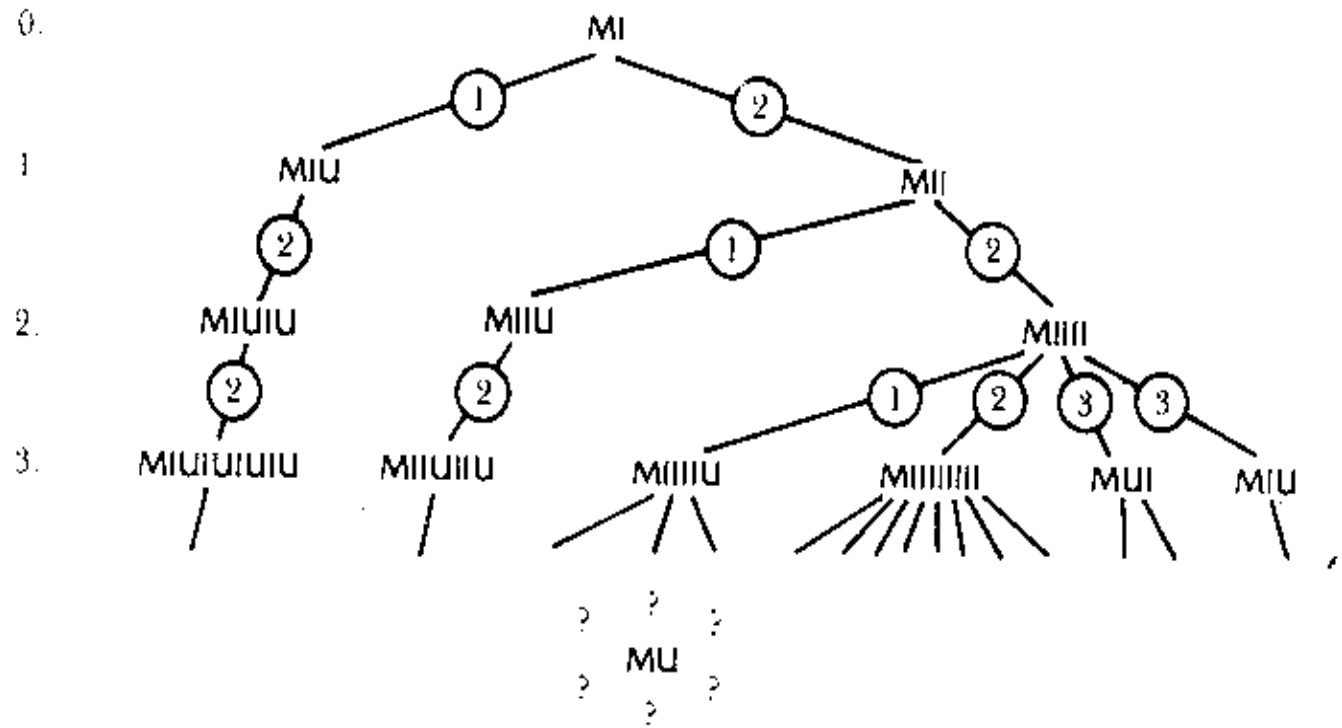


FIGURE 11. A systematically constructed "tree" of all the theorems of the MIU-system. The N th level down contains those theorems whose derivations contain exactly N steps. The encircled numbers tell which rule was employed. Is MU anywhere in this tree?

图27 MIU系统中所有定理的系统结构树

图中的第N级正好包括那些需要N步才能从MI推出的定理。图里的数字则表示所运用的规则。那么MU是否在这棵树上呢?

图28　图案画　　斯科特·E·金（1975）

编委名单

《走向未来》丛书编委会

主　编：包遵信

副主编：[illegible]　唐若昕

编　委：（按姓氏笔划为序）

王岐山　王晓强　王晓鲁　尹蓝天

刘青峰　乐秀成　朱嘉明　朱嘉豪

阮芳赋　何为凌　严家其　陈子伶

易小冶　张　纵　贾新民　陶德荣

翁永曦　黄江南　董秀玉　樊宏业

戴士和

编者献辞

《走向未来》丛书和读者见面了。

她凝聚着我们的心血和期望。

我们期待她能够：展现当代自然科学和社会科学日新月异的面貌；反映人类认识和追求真理的曲折道路；记录这一代人对祖国命运和人类未来的思考。

我们的时代是不寻常的。二十世纪科学技术革命正在迅速而又深刻地改变着人类的社会生活和生存方式。人们迫切地感到，必须严肃认真地对待一个富有挑战性的、千变万化的未来。正是在这种历史关头，中华民族开始了自己悠久历史中又一次真正的复兴。

在艰苦而又富有生命力的改革道路上，我们坚

定了马克思主义的信仰，理解了科学的价值，并逐步深化了对我们时代和民族的认识。今天，我们听从祖国的召唤，热情地投身于实现社会主义现代化的伟大潮流。

马克思有一句名言："思想的闪电一旦真正射入这块没有触动过的人民园地，德国人就会解放成为人。"* 今天，照亮我们民族的思想闪电，就是马克思主义、科学精神和我们民族优秀传统的结合，以及由此开始的创新！

在中国社会科学院青少年研究所领导下，我们编辑了这套《走向未来》丛书。

她力图从世界观高度把握当代科学的最新成就和特点，通过精选、咀嚼、消化了的各门学科的知识，使读者特别是青年读者能从整个人类文明曲折的发展和更迭中，理解中华民族的伟大贡献和历史地位，科学地认识世界发展的趋势，激发对祖国、对民族的热爱和责任感。

她特别注重于科学的思想方法和新兴的边缘学科的介绍和应用；把当前我国自然科学、社会科学，以及文学艺术方面创造性的成果，严肃地介绍给

---

* 《马克思恩格斯选集》，人民出版社，1975年版，第1卷，第15页。

社会，推动自然科学与社会科学的结合。

《丛书》是个新园地，她将自始至终贯彻严肃认真的学风和生动活泼的文风。

《走向未来》丛书从她一开始就受到老一辈共产党人关怀，受到学术界前辈的热情支持。

约四百年前，弗兰西斯·培根在《伟大的复兴》一书序言中，曾经这样谈到书中描述的对象，他"希望人们不要把它看作一种意见，而要看作是一项事业，并相信我们在这里所做的不是为某一宗派或理论奠定基础，而是为人类的福祉和尊严……。"我们怀着真挚的感情，把这段话献给《丛书》的读者，希望广大读者关心她、批评她、帮助她。

让她成为我们共同的事业。

《走向未来》丛书编委会

一九八三年六月于北京

序　言

1979年有一部书轰动了美国，这就是获得普利策大奖的《GEB——一条永恒的金带》。G即哥德尔（Gödel），他是本世纪最伟大的数学家之一。以哥德尔命名的定理是数学理论大厦的高深组成部分，是数理逻辑、人工智能的基石。E是埃舍尔（Escher），当代杰出的画家。B是巴赫(Bach)，最负盛名的古典音乐大师。

本书揭示了数理逻辑、绘画、音乐等领域之间深刻的共同规律，指出了一条永恒的金带把这些表面上大相径庭的领域贯穿在一起，它还构成了奥秘的思维、人工智能和生命遗传机制的基础。

本书作者道格拉斯·霍夫施塔特是一位风华正茂的数学家，是《科学美国人》游戏数学专栏新任

的撰稿人。他是围绕着哥德尔定理来展开全书论述的。

我们知道，哥德尔定理是数学发展史上一座划时代的里程碑。它宣告了在数学发展过程中占主导地位的形式主义流派最终目的的破产。作为严密艰深的数学论文，哥德尔的工作很少为专家圈子之外的人们所了解。但是作为重要而深刻的理论成果，哥德尔理论又具有独特的魅力。这样就更加增添了在这个理论周围的神秘色彩。

为了使中国新一代的青年也能了解这个被视作高深莫测的理论，我们改写了霍夫施塔特的同名原著。原著是一本巨著、堪称是一部“奇书”。作者以生动的语言、广博的知识，从各个角度介绍了哥德尔定理，以漂亮的方式、清晰的思路展开了这个定理的陈述。作者的笔墨洋洋洒洒、纵情驰骋，从巴赫驾驭卡农的精湛技艺谈到埃舍尔令人眩目的版画，忽而又以人们意想不到的方式，把严谨的数学理论与分子遗传神秘的世界联系起来。书中既有各种逻辑符号，也有关于印度数学奇才拉马努贾的风趣描述。读者会不知不觉地被引入一个扑朔迷离的世界，使人们能在各个似乎遥遥相隔的人类活动领域中发现如此惊人的一致性。这就使得读者能够对

于20世纪最伟大的数学成果之一哥德尔定理有一种深刻的印象和切身的感受。

但是全书的论述又是紧扣主题、层层深入的。文章从音乐、美术中的怪圈谈起，继而又谈到逻辑和数学中的怪圈。怪圈可以说是全书的中心概念，也是作者对于哥德尔定理本质的最形象的概括。作者特别强调了赋于形式符号以生命活力的同构，介绍了形式系统及其内部结构，然后又论述了另外一个极为重要的概念——递归。为了说明原始递归与递归的区别，作者形象地用反衬画和互衬画中图形与背景的相互关系来作比喻。在这些重要概念的基础上，作者阐明了哥德尔定理证明过程中最重要的两个思想，即哥德尔编码与自我相关。作者介绍了如何在这个重要思想的基础上构成定理的证明，然后进一步指出，形式系统的这种漏洞是永远无法弥补的。此外，文章还论述了这些思想在理解思维过程、人工智能和生物遗传机制中的重要性，表现了哥德尔理论影响的深度和广度。

原著不仅在内容上精心推敲、而且在编排形式上颇具匠心。原书共有20章，每一章前面都有一段妙趣洋溢的对话。对话中的主角就是古希腊哲学家芝诺在他的疑难中创造的角色，即阿基里斯和乌龟。阿

基里斯是跑得最快的希腊英雄，乌龟则是大家熟悉的爬得极慢的动物。阿基里斯和乌龟的赛跑是众所周知的典故。选用这两个角色便使文章更增添了哲理和悖论的色彩。每一段对话还巧妙地以巴赫杰作《音乐的奉献》中一段贴切的曲子为标题。而对话的内容则以一种隐喻的方式启示了这一章所要阐述的主题。与此同时，作者在书中选用了许多寓意深刻的图画，主要是埃舍尔的作品。于是，埃舍尔的画、巴赫的曲子以及关于歌德尔定理的阐述美妙地编织在一起，组成了一条光灿灿的金带。而这条金带的结尾又极自然地与文章的开头衔接起来，从而构成了一个光彩夺目的金圈。

可惜因为改写篇幅的限制，我们不得不改变了这种结构。那些对话大部分都被删去了，只保留了少数几段并大大加以压缩，然后编入各章之中。读者可以从“奇妙的唱机和唱片”、“历险记”、“没有结束的唱机之战”这些章节来领略一点原书的风味。

为了使哥德尔定理的阐述更容易为读者理解，全书的章节也重新作了编排，内容则作了相应的改写。当然，作者的那些重要思想，我还是尽力想把它忠实地表现出来。

作者反复阐述的怪圈、生动描述的同构以及用图形和背景来帮助人们理解那些极为枯燥的数学概念，实在是令人拍案叫绝。

尽管这本书的例子生动、思想新颖，但是要想不费功夫或者较轻松地读通它仍然是做不到的。中国与西方的文化背景不同。一般人对于巴赫的音乐、埃舍尔的画是极为陌生的，而这里介绍的数学理论又往往会使人望而生畏。因此不能指望大部分人读一遍就能看懂全书的内容。

但是这本书的文字和图画富有机智和启发性。读者不妨先细细品味一番前面的图画，参看图下的简单说明，这将是一种很微妙的享受。我们也不妨把阅读此书看成是在挖掘一个价值很高的矿。需要你付出辛勤的劳动，但会使你有丰硕的收获。也可以反复地看，每看一遍会有新的体会、新的理解、新的启迪。有时最好还能拿起笔来演算一番。当然最重要的还是能够认真地进行思考。如果读者对于书中的思想有所领悟，无疑会影响到他的思维方式。

哥德尔的理论改变了数学发展的过程，触动了人类思维的深层结构，它又渗透到音乐、艺术、生物、计算机和人工智能等领域。如果读者能够通过这本书汲取一点养料，或者从一幅画一段叙述中有

所领悟，那就令人欣慰了。

乐秀成

1983年3月于北京

目　录

序言…… 1

1　怪圈…… 1

1.1　绘画与音乐中的怪圈…… 3

1.2　怪圈与悖论…… 7

1.3　数学和思维中的怪圈…… 14

2　同构…… 19

2.1　奇妙的唱机与唱片…… 21

2.2　符号的意义…… 28

2.3　破译…… 32

2.4　音乐和绘画中的同构…… 40

3 形式系统…… 43
3.1 从一个数学难题谈起…… 45
3.2 一致性和完备性…… 51
3.3 形式系统的结构…… 54
3.4 命题演算系统…… 57
3.5 形式数论系统…… 63

4 递归…… 71
4.1 阿基里斯和乌龟历险记…… 73
4.2 形形色色的递归…… 78
4.3 图形和背景…… 88

5 哥德尔定理…… 97
5.1 原始递归与递归…… 99
5.2 哥德尔编码与自我反省…… 107
5.3 无法弥补的漏洞…… 116

6 大脑与思维…… 121
6.1 描述的层次…… 123
6.2 大脑…… 132
6.3 塔式基—丘奇—图林定理…… 139
6.4 思维…… 149

7 人工智能…… 155
7.1 回顾…… 157

7.2 人工设计的游戏…… 165
7.3 概念框架与模式识别…… 171
8 自我相关与自我复制…… 177
8.1 遗传机制的形式化…… 179
8.2 遗传机制的支持系统…… 187
8.3 没有结束的唱机之战…… 192
8.4 自相缠绕的层次…… 196
9 结束语…… 201

1

怪　圏

1.1 绘画与音乐中的怪圈

我第一次知道埃舍尔的名字是在十多年前。那时我津津有味地看着诺贝尔物理学奖金的获得者杨振宁博士所著的小册子《基本粒子发现简史》。我特别注意到杨振宁先生在前言中对埃舍尔先生允许他采用《骑士图》表示深深的谢意。我被这张图深深地吸引住了，因为埃舍尔以优美的图形及其镜象巧妙地表现了对称性的原理。这些原理在物理学的世界中起着极为重要的作用。可惜，迄今为止，中国人一般还不熟悉埃舍尔的作品，但是在西方他

是一位别具一格、极有影响的画家。埃舍尔创造了一系列富有智慧的图画，其中有许多画体现了奇妙的悖论、错觉或者双重的含义。因此，在埃舍尔作品的崇拜者中间有许多数学家也就不足为怪了。当我们慢慢欣赏埃舍尔的画并在其中发现那些美妙的数学原理时，那是一种多么愉快的享受啊！

埃舍尔的画往往表现了一些很深刻的思想，怪圈就是其中最常见的一种。我们先来看看那幅奇怪的版画《瀑布》（图1）。在画面的中央，瀑布倾泻而下，水花四起，还推动了水轮。汇集到水池中的水则顺着水渠哗哗地流去，一级一级地下降。突然水又流到了瀑布口！真是不可思议，可是在画面上却表现得明明白白。我们只能把这种周而复始的圈称作“怪圈”。

我们再看那幅《上升与下降》（图2）。在这冰冷阴森的教堂里，僧侣们排成两队往前走。其中一队总是沿着楼梯往上走，另一队总是往下走。可是他们走的却是同样的楼梯，并且不断地回到原来出发的地方。真是妙不可言。这又是一个怪圈。所谓怪圈就是指这样一种现象，我们在某一个等级系统中逐步上升（或者下降），结果却意外地发现又回到了原来开始的地方。有时我就用“缠绕的层

次”来描述其中有怪圈的体系。怪圈还要在本书中一再出现，忽上忽下，时隐时显，希望读者能够细细地体会它的内在含义。

有趣的是在音乐中也有这样奇妙的怪圈。为此我们先介绍一下近代西方音乐的鼻祖，被誉为“音乐之父”的J·S·巴赫。巴赫是颇负盛名的钢琴家、风琴家和作曲家。当时的普鲁士国王弗里德利希是他的崇拜者。他把62岁的巴赫邀请到自己的宫廷来，并向他展示了自己收藏的钢琴。这种钢琴在当时还是问世不久的珍品。

巴赫在每一架钢琴上进行即兴表演，使得弗里德利希大为倾倒。当巴赫返回莱比锡后，他收到了国王自己创作的一部分乐谱。巴赫在这些乐谱的基础上写成了举世闻名的主题乐曲《音乐的奉献》，并把它奉献给弗里德利希国王。巴赫在这部作品中充分发挥了形式上的技巧。有趣的是，在这种登峰造极的技巧中就包含着怪圈。

为了理解音乐创作中的怪圈，我们必须先谈谈“卡农”。卡农是音乐家们熟悉的，它就是重复地演奏同一主题。最简单的方式就是用不同的音部重复演奏，每个音部都比前一个音部延迟一段时间。大部分乐曲的主题与这种演奏方法是不协调的。适

合多音部的主题必须使每个音符具有双重(或多重)的功能，它既是主题中的一部分，又必须与其他音部保持和谐。

在《音乐的奉献》中，用一种特殊的卡农技巧构成了怪圈。它由三个音部组成。当最高音部演奏主题时，其余两个音部提供卡农式的协奏。这种卡农最大的特点就是神不知鬼不觉地进行变调，使得结尾最后能很平滑地过渡到开头。这种首尾相接的变调使听众有一种不断增调的感觉。在转了几圈之后，听众感到已经离开原来的调很远了。可是奇妙的是通过这样的变调又能回到原来的调上。这就是音乐中的怪圈。我们可以体会巴赫的创作意图，无疑他有这样一种想法，采用这种方法可以使升调的过程无限地进行下去。因此他在乐谱上专门注上了“陛下的荣耀也随着变调而增高”。我们不妨把这种卡农称为“无限升高的卡农”。

如果我们把“无限升高的卡农”与埃舍尔的画《瀑布》以及《上升与下降》作个比较，就可以发现两者的相似性是极为明显的。巴赫和埃舍尔采用不同的艺术形式：音乐和美术，却表现了同样的思想：怪圈。

埃舍尔用绘画表现的怪圈有许多不同的形式，

有松弛的也有紧凑的。在《上升与下降》中，这种怪圈是比较松弛的，僧侣们要经过许多级才能返回原处。而在“瀑布”中怪圈就要紧凑一些，它总共只有6级。你也许已经想到了，这里“级”的计算有含糊不清之处。例如我们可以把《上升与下降》的系统算成是45级（按台级算）也可以算成是4级（按楼梯算）。这种模棱两可性不仅表现在埃舍尔的画的怪圈上，也表现在其他形式怪圈的系统中。更紧凑的怪圈可以在《画画的双手》（图22）这幅画中看到。而最紧凑的怪圈要数《画廊》（画23）了。这幅画中之画包含着自身。我们可以说画廊中的一幅画包含着它自身，也可以说这个城市包含着它自身。

怪圈的内在含义也是在有限中包含无限的概念。它不仅仅是一个圈，而且是以一种有限的方式来体现无限的过程。这在埃舍尔的著名作品《变形》（图4）中表现得极为明显。我们从作品的某点出发，随着画面的逐级变化而向前走去，走着走着却突然回到了原来出发的地方。

1.2 怪圈与悖论

在巴赫和埃舍尔创造的这些怪圈中，存在着无

限与有限的矛盾，荒唐与真实的对比，往往会给人以强烈的悖论感。这种直觉表明，在怪圈中包含着深刻的数学原理。事实也确实如此。就在我们生活的这个世纪里，有一个影响深远、与之呼应的重大数学发现，这就是哥德尔在数学系统中发现了怪圈。这种怪圈可以说是起源于一个古老的逻辑悖论。它在历史上被称为爱皮梅尼特悖论。

爱皮梅尼特是一个克里特岛人。他说："所有的克里特岛人都撒谎。"假如他说的话是对的，那么作为克里特岛人的爱皮梅尼特就是在撒谎，那么他的话就是错的。反之，假设他的话不对，那么作为克里特岛人的爱皮梅尼特就没有撒谎，他的话就是对的。无论采用哪一种假设，都是无法自圆其说的。我们也可以把这个悖论表述成更为简洁的形式。这就是"我说的这句话是错的"。这是和《画廊》一样的单级怪圈。因为这个句子中的"话"可以指这个句子本身。这就是说一个句子在描述这个句子本身。

以后人们又发现了许多其他形式的悖论。尤其是在本世纪初，随着集合论与数理逻辑的发展，在数学和逻辑中发现了许多悖论。其中最著名的有康托尔悖论和罗素悖论。在这些悖论中好象都有一个

共同的“罪犯”，这就是自我相关，或者就是我们所说的“怪圈”。罗素悖论用形象的语言来描述，就是有一位理发师声称，他给所有不给自己理发的人理发。那么这个人是否给自己理发呢？如果他给自己理发，就违背了自己的声明。如果他不给自己理发，也没有兑现自己的诺言。用集合论的述语来说，罗素悖论就是定义这样一个集合A，它由所有不属于A的元素a组成。那么A是否属于它本身呢？如果A不属于A，那么按照集合A的定义，它就属于A。如果A属于A，那么按照定义它就是不属于A的元素。显然在罗素悖论中，最关键的地方就是假定一个集可以自己属于自己。这就是自我相关。因此，要想排除悖论很自然就会想到，要防止自我相关和造成自我相关的条件的出现。

罗素和怀特海就是在这种思想的指导下写出《数学原理》的。他们竭力想把“怪圈”从逻辑、集合论以及数论中驱除出去。他们的基本思想是把集合分成各种等级。最低一级的集合只能以那些“对象”而不能以其他集合为元素。较高一级的集合只能以对象或者更低级的集合为元素。这样每个集合都被安置在某一个等级上，也就排除了一个集合以自己为元素的可能性。

如果我们把所有的集合分成两类。第一类集合不能以自己为元素，也就是说自己不能属于自己，我们称为r型。第二类集合可以以自己为元素，我们称为s型。那么在《数学原理》规定的系统中就只有r型的集合。这样进行分级，可以使集合论中不再出现悖论。但是付出的代价是，必须引进人为的分级和禁止生成某种类型的集合。例如所有集合的集合，或者所有不属于A的集合，也就是我们所说的s型集合。

这种理论可以用来对付罗素悖论，但是无法对付爱皮梅尼特悖论。因为用同样的方法对付爱皮梅尼特悖论就要对语言进行分级。于是就有所谓的对象语言、有描述语言的元语言，还有描述元语言的元元语言，等等。

我们模仿《画画的双手》把上述悖论写成两句的形式：

下面这句话是错的。

上面这句话是对的。

如果按照分级理论的规定，上面这句话是描述下面这句话的，因此它属于更高一级。但是按照同样的道理，下面这句话应该比上面这句话属于更高一级。因此这两句话不能同时满足分级理论的要

求，也就是说它们不能同时有意义。

如果说把集合进行分级的理论还是貌似有理的，那么把语言进行分级就是十分荒唐的。当我们谈论各种事物时，是决不会意识到自己在不同的层次之间上窜下跳的。例如“我在这本书中评论了分级的理论”。这是很普通的一句话。但是在严格分级的语言系统中它就要受到双重的禁止。首先，在这句话中谈到了“这本书”，它只能在“元书”所属的层次中出现，而不能在“这本书”所属的层次中出现。其次，这句话居然谈到了我，这是我无论如何不允许谈论的！这个例子清楚地说明了，如果把严格分级的理论引进我们所熟悉的日常语言中来，将是多么荒谬。由此可见，如果用这种方式来弥补由于悖论出现而造成的缺陷，即排除任何形式的自我相关，实际上是走过头了。它把美妙的语言结构变成了没有价值的僵尸。

也许有人会辩解说，这种分级的理论可以适用于形式语言而不是我们用的日常语言。但是这恰好说明了，这种理论只有在我们人为拼凑的系统中，为了避免悖论才有用处。而为此作出的牺牲，却是在这种人为的系统中加上生硬的限制。这样的理论系统虽然是一致的，却是乏味的、令人生厌的。你

一定会感觉到这里有什么地方不对头了。

不过还应该指出的是，在本世纪初，数学家们并没有把语言中无法排除怪圈看得那么严重。他们至少可以这样安慰自己，语言是不严格的，而数学却是严密无隙的，只要能在数学理论中排除怪圈就可以了。于是数学家们为自己确立了这样的目标，建立一座形式系统的大厦，这座大厦的基石是一些公理，然后严格按照形式逻辑推导出系统中的每一个定理，而整个系统是完全确定的，不会相互矛盾的。在这样的系统中可以排除任何悖论的出现。这就是《数学原理》的目标，它要用逻辑来推导出所有的数学成果而又不会产生矛盾！这种目标确实是激动人心的，但是当时谁也无法肯定，采用罗素和怀特海所提供的方法，能否囊括所有的数学成果，也无法知道，这种方法能否永远保持一致，即完全排除悖论的出现。

杰出的德国数学家（也是研究数学基础的元数学家）希尔伯特沿着罗素开辟的道路勇往直前。他向数学界提出了一个明确的任务：严格地按照罗素和怀特海所描述的方法，证明《数学原理》所定义的系统既是一致的（无矛盾）又是完备的（该系统的理论框架中容纳了每个正确的数论命题）。这就

是数学史上著名的希尔伯特纲领。当希尔伯特提出这个纲领后，就有人尖刻地批评这种说法是一种循环，你怎么能根据自己的推理方法来判定这种推理方法的正确性呢？这似乎是要抓住自己的头发把自己举起来。（我们看来是无法摆脱这种该死的怪圈了。）

希尔伯特意识到这种困难性。他进一步加以说明，这种关于一致性和完备性的证明只能依赖于“有限”的推理步骤。这个目标曾经在本世纪的前30年中使许多伟大的数学家绞尽脑汁。但是到了1931年，哥德尔发表了他的论文《论〈数学原理〉中形式上不可判定的命题及其有关系统Ⅰ》。这篇论文彻底推翻了希尔伯特纲领，因为它指出了没有一种公理系统可以导出数论中所有的真实命题，除非这种系统是不一致的，即存在着互相矛盾的悖论。因此企图证明《数学原理》所示系统的一致性是徒劳的。如果能够找到一种证明，仅仅使用《数学原理》中的方法，那就会得到哥德尔定理最神秘的结论：《数学原理》本身是不一致的！于是怪圈成了逻辑和数学中无法驱除的幽灵。

1.3 数学和思维中的怪圈

哥德尔的研究成果使得希尔伯特纲领陷于破产。从此以后，数学家意识到，在严密的数学理论体系中也存在着漏洞。这种意识从某种意义上讲是痛苦的。过去被人们视为神圣的、万能的数学从理想天国的宝座上跌落到了尘世人间，它不再是完美无缺的。对于哥德尔定理的严格论证向来只是少数专家问津的领地。但是就和人类思想的其他伟大成果一样，哥德尔的基本思想是既深刻又清晰明了的。

哥德尔的重要思想就是用数学推理来探索数学推理本身。这种使数学“反省”自己的概念被证明是非常有效的。它的主要成果就是哥德尔不完备定理。如果我们把这个定理比作珍珠，那么哥德尔的证明方法就是产生这颗明珠的贝母。珍珠以它的光彩和纯洁引人瞩目；贝母却是一种复杂的生物体，它的内部结构产生了这颗奇妙纯真的瑰宝。

哥德尔定理指出，数论的无矛盾公理化的所有陈述中必定包含着不可判定的命题。

从这颗珍珠本身很难直接看出其中的怪圈，因

为它被掩藏在自己的贝母——它的证明过程中了。这个证明是以自我相关的形式语言写成的。哥德尔意识到可以把自我相关进行表述的思想运用到数论体系中来，他创造了这种表述的方式，于是他就克服了主要的障碍。

为了阐述哥德尔思想，需要先说明一下它所涉及的数论系统。数学家们可以证明，非欧几何公理系统的无矛盾性与欧几里得几何公理系统的无矛盾性是等价的。而欧几里得几何公理系统的无矛盾性又与算术公理系统即数论系统的无矛盾性等价。因此我们可以把数论看成数学理论大厦中最基本的组成部分。只要证明，在数论的形式系统中存在着不可判定的命题，那么也就证明了，在数学理论中无法把怪圈排除出去。但是我们知道，一般来讲，整数既不是一种陈述，也不是关于它们性质的陈述。因此数论中的陈述也不是关于数论陈述的陈述。那么如何才能使数论反省自己呢？

哥德尔意识到，要做到这一点，首先要把命题的陈述及命题的推理过程数字化。只有用数字以某种方式代替命题的陈述，命题的推理过程才能成为数字的运算，这种运算过程才能数字化。只有这样，形式数论系统中的陈述才能成为关于数论陈述的陈

述。他终于找到了这种方法，这就是通常称为哥德尔编码的方法。这是一种逻辑命题的各种形式符号与数之间的同构关系。哥德尔编码是整个理论的结构核心。运用这种技巧，数论中的陈述就可以从两种不同的层次来理解：既是数论中的陈述，又是关于数论陈述的陈述。

有了哥德尔编码，我们就可以设法把爱皮梅尼特悖论搬到数论的形式体系中来。不过，需要说明一点，他最后移植过来的并不是这样的定理，“数论中的这个命题是错的”，而是说“数论中的这个命题是无法证明的”。

如果我们把爱皮梅尼特悖论看成是形式逻辑系统中无法驱除的怪圈，那么哥德尔编码就建立了形式逻辑系统中命题与数论中某些数的同构。他还进一步实现了在形式数论系统中的自我反省。也就是说，他在形式数论系统中构造了与上述悖论同构的怪圈。

哥德尔理论指出，由于自我相关的怪圈存在，人们面临着二择一的两难境地。要么在逻辑思维中可以是不一致的；要么导致另一个意想不到的结果，我们无法用逻辑去证明所有看来是用逻辑提出的问题，这就是不可判定性。数学家们接受了后一种选

择，因为严格的数学理论如果允许不一致性就会导致数学大厦整体的崩溃。

当然，哥德尔定理对任何一致的公理系统都是适用的，因此这个成果对于逻辑学家、数学家以及对于数学基础感兴趣的哲学家都是一个巨大的冲击。它宣告，无论怎样复杂的形式系统都无法表现整数的完备性。这在当时看来，简直是祸从天降！

当哥德尔的论文问世时，世界正处于发展电子计算机的边缘。不过计算机的发展历史可以追溯到巴斯卡、莱布尼兹和贝比奇的时代。与哥德尔的数学成果相对应，图林在计算机科学理论中指出了，即使可以设想的最有效的计算机也存在着无法弥补的漏洞。后来人们发现这个结论与哥德尔定理是等价的。于是哥德尔理论的影响便越出数学的疆域而扩展到人工智能及思维的研究。

到了20世纪50年代，机器智能看来迫在眉睫了。电子计算机的功能迅速扩大，许多以前被人们认为只有人的思维才能完成的智力活动，逐渐都由电子计算机轻而易举地应付过去了。但是，随着每一道旧的障碍被克服，又出现了新的障碍。谁也不知道人的智力行为与非智力行为的确切界线究竟在哪里。也可能认为存在这种明确界线的想法本身就

是愚蠢的。

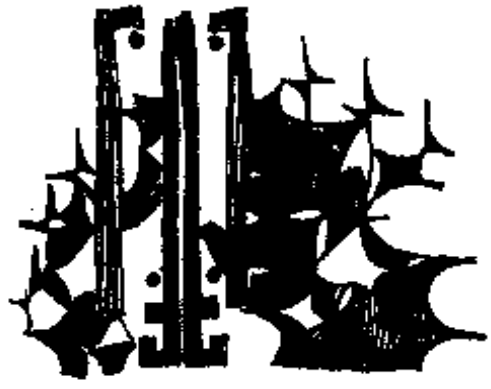
人工智能的研究最感兴趣的就是把一大堆严格形式化的规则搭配在一起，告诉不灵活的机器如何变得灵活起来。但是什么样的“规则”才能控制智力行为呢？这些规则一定可以分成不同的等级。有许多规则是“简单明白的”。还有一些是修改简单规则的元规则。然后还有元元规则。人工智能的灵活性就来自大量的不同规则和这些规则所划分的不同层次。这些层次的自相缠绕就是人工智能中的怪圈。这直接或者间接地关系到人工智能的核心。

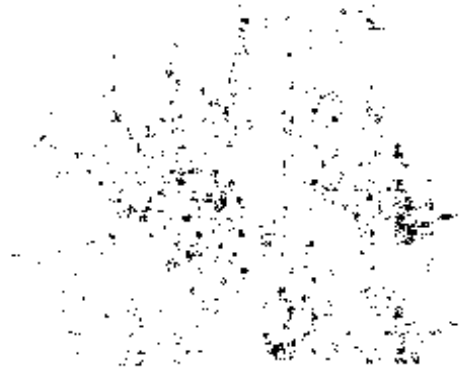
哥德尔定理和图林定理并不意味着机器是不可思维的。事实上，现在还无法说清楚人的思维究竟是一种什么过程。人们对于自己的大脑研究越深入，对于思维过程了解得越多，就越感到原来的想法是多么错误。

大脑的活动是以神经细胞为基础的。但是仅仅在这个层次上是无法理解大脑思维活动的。人工智能的研究启示我们，在思维过程中存在着错综复杂的层次，这些层次的自相缠绕很可能在思维过程中起着关键性的作用。其实人的思维早就开始探索思维本身，这也是一个绝妙的怪圈！

2

同　　构





2.1 奇妙的唱机与唱片

为了说明怪圈与哥德尔定理之间的关系，先让我们来讲一段小故事。

（阿基里斯去访问乌龟，并在他家消磨时间。）

阿：上帝啊！你的收藏品可真多。你收集了这么多唱片，那你究竟喜欢什么样的唱片呢？

龟：我认为巴赫的作品最棒。不过我最近感兴趣的

却是一种特殊的音乐。我把它称作“粉碎唱机的音乐。”

阿：这可真是一种古怪的音乐。难道是你举着大锤，按照贝多芬《惠灵顿的胜利》的节奏把唱机一个一个地砸碎？

龟：可不是这么回事。懂得这种音乐的人并不多。这要从我的朋友蟹来说起。有一天他来我这儿作客。他刚刚买了一架新唱机。按照店主的说法，它能重演任何声音。也就是说这是一架完备的唱机。

阿：你肯定是不相信这一点的。

龟：后来我就去回访他，并且带去一张我自己创作的唱片。唱片的曲名就叫做：“我不能在唱机上演奏。”我建议他和我一起来欣赏这张唱片。于是他就打开唱机把这张唱片放进去了。不幸的是，刚奏出几个音符，唱机就开始抖动起来、越抖越厉害，最后只听见“啪”的一声，唱机裂得粉碎。不用说这张唱片也跟着报销了。

阿：真倒霉，可是店主不是吹嘘这是一架完备的唱机吗？

龟：确实如此。阿基里斯，难道你也会和蟹一样天真，相信店主告诉你的一切吗？

阿：我想这是因为店主在吹牛的缘故。

龟：其实我在回访蟹之前就去过出售唱机的那家商店。我向他索取了设计说明书，分析了它的结构，并且发现确实有这样一组声音，如果它在唱机附近作响，就可以使唱机振荡，乃至于碎裂。

阿：你这个恶毒的家伙！不用细说我也明白了。你录下的真是这组声音，还假惺惺地把它当作礼物去送给蟹。

龟：你倒是够机灵的。不过事情并未因此了结。蟹并不相信他的唱机是有缺陷的。于是他买了一架更加昂贵的唱机。店主则向他许诺，如果他能发现一组在这架唱机上无法重演的声音，就包赔两倍的钱。于是蟹就兴致勃勃地来找我。而我也很有兴趣再去看看。

阿：我敢打赌，你一定又按照新唱机的结构泡制了一张新的唱片："我不能在唱机2上演奏。"

龟：你的思维很清楚，完全领会了问题的精神实质。当然，完全可以料想到，这架唱机又被震得粉身碎骨了。

阿：我倒有一个主意。他可以买一架低保真度的唱机。这样就不会再重演使它自己毁坏的那组声音了。

龟：可是这样一来就违背了原来的宗旨——可以重演任何声音。

阿：我现在明白问题的两难性究竟在哪里了。这就是说任何唱机其实都是有缺陷的。

龟：我不明白你为什么要把这叫做缺陷。问题的实质在于，你要唱机去做它根本办不到的事情。不过我的朋友蟹并不死心。他又自己设计了一架“奥米伽唱机”。这种唱机带有一架电视摄象机，能在唱片演奏之前先把它审视一番。它和微型计算机联结在一起，可以立即判定这组声音的性质。通过精密的计算，它可以知道这组声音对于唱机所产生的效果。如果唱机会受到破坏，它就可以通过一个内部装置将唱机的各部分重新组装，从而改变它的内部结构再来演奏唱片。

阿：这下好了，你也没有办法了吧。

龟：瞧你这副得意的劲儿，如果你懂得哥德尔定理，就不会这样得意了。

这个故事有趣地说明了哥德尔定理的实质。存在着不能演奏的唱片就相当于存在着不可判定的定理。之所以存在着不能演奏的唱片就是因为唱机和声音之间必定存在着自我相关的怪圈。我们仔细分

析一下这个故事，就可以看到，这个故事中的对话具有双重的意义——明显的和隐含的。唱片的纹沟也有两层意思。第一层意思，这是一种音乐。纹沟通过唱机成为空气的振动。第二层意思则是空气的振动反过来感应出唱机的振动。这第二层意思依赖于两个同构的链：

（1）纹沟模式与空气振动之间的同构。

（2）空气振动与唱机振动之间的同构。

这里出现的“同构”是本章所要阐述的中心概念，也是理解哥德尔定理的关键之一。顾名思义，同构就是构造相同。

我们知道，构造是事物之间的一种关系。数学系统的构造则表现为一些运算法则。可以用数学语言对于两个数学系统的同构进行精确的定义。两个同构的数学系统可以看成是等价的。有趣的是数学家发现两种结构之间的同构关系，往往象从天上掉下来一样突然。实际上，两种结构之间同构关系的发现在知识的进步中具有重要的意义。可以说，通过对同构的意识才使形式产生了意义。同构是一支神笛，它使无声无息的木偶获得了生命，成了活蹦乱跳的姑娘。采取更加直截了当的说法，同构就是一种保持信息的交换。

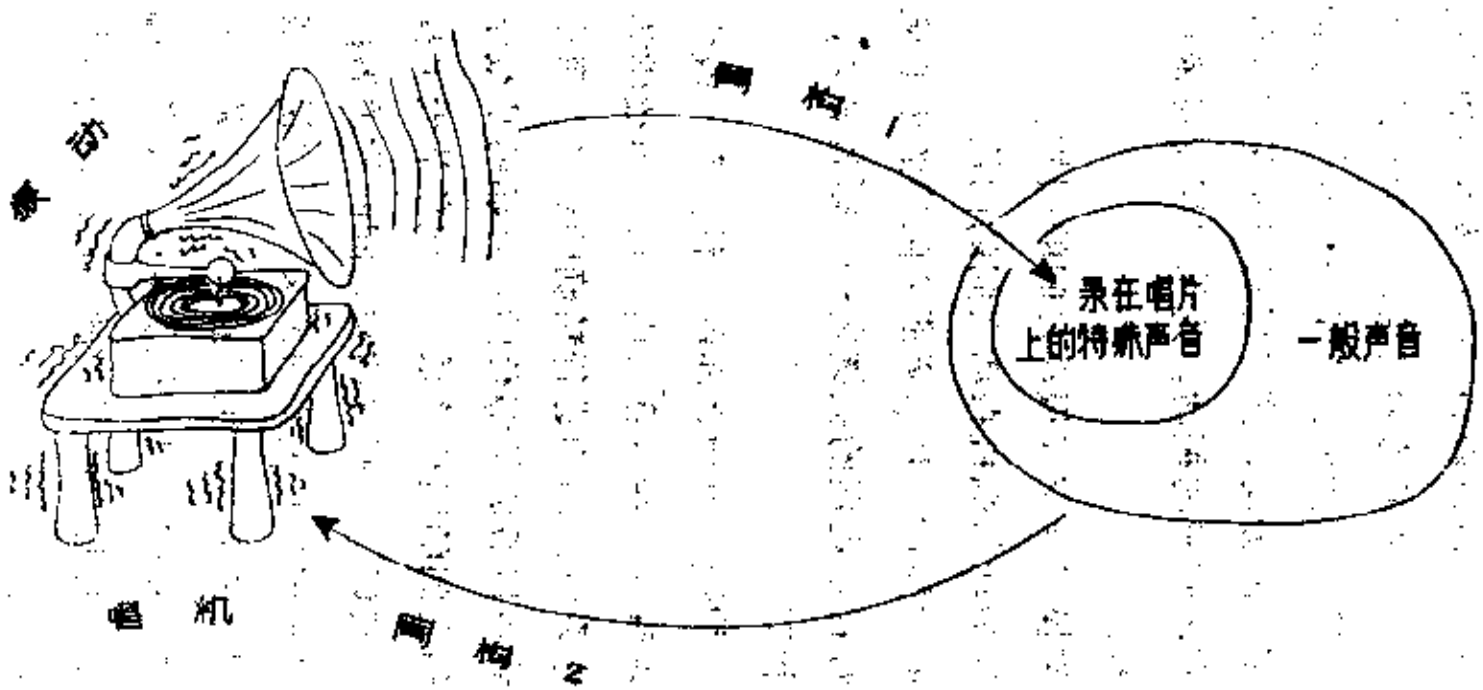


图29　哥德尔定理的基本原理

两个同构产生自食其果的效应。第一个同构从唱片的纹构到唱机发出的声音。第二个同构从唱机的声音到唱机自身的振动。但要注意，第二个同构并不依赖于第一个同构。因为任何声音都会使唱机振动，而不仅仅限于唱机所发出的声音。这两个同构形成了怪圈，使得任何唱机都存在着不能在上面演奏的唱片。

回到我们所讲的故事中来，我们可以得出这样一组对应关系：

唱机⟺数论的公理系统

低保真度唱机⟺“弱”公理系统

高保真度唱机⟺“强”公理系统

“完备”唱机⟺完备的数论系统

唱片⟺形式系统中的串

可演奏唱片⟺公理系统中的定理

无法演奏的唱片⟺公理系统中的非定理

声音⟺数论的真实陈述

重演的声音⟺系统中加以解释的定理

无法重演的声音⟺真实的陈述而不是定理

唱片标题⟺哥德尔串的隐含意义：

“我无法在唱机x上演奏。”⟺“我无法在形式系统x中推导出来”。

这些对应关系并不是上述故事与哥德尔定理同构的全部内容，只是其中的主要部分。那么为什么会有这种对应关系呢？就是因为同构。在这个故事中，为什么会有无法演奏的唱片，这是容易理解的。唱片、声音、唱机之间的同构关系是形象的、明白易懂的。在形式数论中也存在着同构，要将一种推理过程形式化、严格化，都离不开同构。正是这些同构构成了不可避免的怪圈。因此我们还要从同构的本质谈起。

2.2 符号的意义

我们已经看到，在唱片与音乐之间存在着同构。虽然这两个系统性质完全不同，但是唱片纹沟的深和浅可以和音乐曲子中声音的强和弱一一对应起来。由此可见，同构本质上是一种映射，通过这种映射，一个系统的结构可以用另外一个系统表现出来。明白了这一点，我们也就可以知道，为什么同构在哥德尔理论中有举足轻重的地位。哥德尔理论要用数学推理来研究数学推理过程，这是借助于逻辑符号与数字之间的同构关系。这种同构关系使得数字具有了意义。一般来讲，形式符号容易给人一种错觉，好象它是人类意志的自由创造，可以和现实的世界毫不相关。然而同样的事实是，那些和我们关系密切的形式符号，如词汇、数字、逻辑符号，都是人类文化进化过程的产物。它们与现实世界有着密切的联系。这座联系的桥梁就是同构。

我们来看这样一个例子。假定有一个形式系统只有3种不同的符号：

p q -

我们把它称为pq系统。系统中的串都由这3种符

号组成。在这个系统中还规定了这样一些规则和公理：

公理：xp－qx－

其中的x只由连字符号－组成。不过x所代表的连字符号串必须是相同的。

规则：　如果x、y、z都代表连字符号的串，已知xpyqz为一个定理，那么xpy－qz－也是定理。

这里的p q 系统是一种典型的形式系统。所谓形式系统是由一组定理及其变换规则构成的系统。形式系统的定理是由形式符号组合而成的串，这些串可以按照规定的法则进行运算而生成新的串。随意规定的串就是公理。按照规则由公理或定理生成的串就是定理。所谓形式系统的推理，就是严格按照形式系统的规则逐行生成新的定理，直到我们所需要的串出现为止。这个过程也叫做形式证明。

粗一看，这个p q 系统只是一些生疏的、莫名其妙的符号和枯燥乏味的规则。但是如果我们运用一下方才谈到的同构，那么就要刮目相看了。

也许有些细心的读者已经注意到了，pq 系统的定理具有加法的含义。我们可以把p看 成加法，把q看成等于。于是看来毫无意义的串

－－p－－－q－－－－－

就成了我们熟悉的算式 2 + 3 = 5 。

在这个例子中，我们得到了一个很好的同构原型。这两个同构系统之间的对应关系为：

p⟺+

q⟺=

－⟺1

－－⟺2

等等

有时我们也把这种对应关系称为翻译。这是一种低层次上的对应。还有一种高层次上的对应，就是形式系统的定理与实际陈述之间的对应。如果你对一个形式系统一无所知，要想了解这些形式符号的意义，就要寻找一种有意义的方式来翻译。也就是说，寻找一种可以产生意义的同构关系。这就象破译密码一样。

不过对于一个形式系统来讲，这种意义只是被动的，而不是主动的。我们可以通过翻译来了解形式系统定理的意义，但是却不能根据意义去创造这些定理。要创造形式系统的定理，只能严格按照形式系统本身的规则。

对于同样的pq系统，我们也可以用一种完全不同的方式来翻译。

p⟺等于

q⟺从……减去

-⟺1

--⟺2

等等

于是前面的串就同构于“2 等于从 5 减去 3”。这意味着，这种形式系统所反映的现实世界并不是单一的。不过这些不同的方面是相互同构的。对于同样的形式符号可以赋于不同的意义，这是一种非常重要的现象。

采用同构的方法，我们可以对一个未知的形式系统中的某个串作出一种解释。但是我们怎么能知道这种解释是合理的呢？这就需要用同样的解释去检验这个形式系统中的其他定理。如果对于系统中每一个定理的解释都是有意义的，那么这种解释就是合理的。如果一个形式系统中的定理是有限的，那么这个检验过程总是有限的。但是我们怎么能够知道一种解释对于一个有无限定理的形式系统也是合理的呢？这个问题就涉及到无限和递归的概念，我们将在后面讨论这些概念。

同构赋于形式符号以意义。这也意味着形式符号可以把握现实世界。然而形式系统真的具有这种

能力吗？或者说，这种能力究竟有多大？埃舍尔有一幅发人深省的画《解放》（图5）。从画面中我们可以看到规则图形与不规则图形（飞鸟）之间的奇妙对比，在两者之间有一个花哨的过渡区域。如果把现实世界比作自由飞翔的鸟儿，把形式系统比作规则的图形，那么这些鸟儿真的能受形式系统规则的约束？是否在现实世界与形式系统之间也存在一个奇妙的过渡区域呢？

2.3 破 译

运用同构和翻译机制可以使我们理解形式系统中符号的意义。我们可以用这种方式去揭示某种结构的内部信息。但是并不是所有结构的内部信息都能用这种方式揭示出来的。这一点在生物遗传和生命过程中十分重要。

按照一般的说法，遗传的信息是贮存在脱氧核糖核酸DNA的双螺旋体中。从DNA分子（基因型）转变成生物体（表现型）是通过一系列非常复杂的过程，其中包括各种蛋白质的制造，DNA的复制、细胞的复制、各种细胞的分化等等。

一般可以把生物体的结构归因于DNA的结构，

归因于这种结构所携带的信息。奥斯瓦德·艾弗里在1946年指导进行的实验首先提供了这方面的证据。他的实验表明，只有DNA才能传递可以遗传的性质。人们可以改变生物体中某些分子的性质、例如某些蛋白质的性质，可以使一些生物体产生很明显的特征。但是这些特征是不能传给下一代的。只有改变了DNA分子，改变的性质才能遗传给后代。改变遗传信息是构造一种新生物的唯一途径。

因此，看来人们被迫要接受这样的观念，在DNA的结构中包含了表现型结构的信息。这就是说，两者之间有着同构关系。但是这种同构并不是一般数学意义上的同构，而是一种奇异的同构。我们之所以要区分“基因型”和“表现型”并不是多此一举。一般意义上的同构简单地将一个系统的结构映射到另一个系统中。唱片与音乐之间的同构就是这样。但是DNA基因型结构与生物体表现型结构之间的同构就完全不一样了。要使这种同构在实际中实现，需要一种极为复杂的机制。如果你想在自己基因的DNA上找到哪一部分与你的鼻子或者指纹的形状有关，那是不可能的。这就好象你要在一段音乐中找到表达激情的音符一样可笑。因为激情这种“意义”是通过较高的层次来表现的，是通

过这段音乐中较大的“块”而不是单个的音符来表现的，是通过一组音符以及它们之间的相互关系来表现的。

同样，“基因的意义”——关于表现型的信息——是包含在DNA的整体中的。显然，现在还没有人能理解这种语言。因为把基因型转换成表现型的机制远比基因型本身复杂得多。我们可以把基因看成一种开关，它能够触发这些机制来达到转换的目的。DNA的一部分能够引发蛋白质的制造。这些蛋白质则能够引发几百种生化反应。这些反应又引起复制DNA本身的操作……这样多级引发的最后结果才是表现型的生物体。也许有人会说，这种表现型揭示了潜在地包含在DNA中的信息。但是恐怕不会有人会认为，自动电唱机的喇叭所播出的乐曲“揭示”了包含在一对键纽里的信息。因为这对键纽只是引发了自动电唱机中与信息有关的机制。但是我们完全有理由说，用唱片演奏出来的乐曲确实揭示了包含在唱片纹沟里的信息。这有几方面的理由：

（1）看来乐曲并不是隐含在唱机的机制中。

（2）可以以任意的精确度使输入（唱片）与输出（乐曲）之间匹配起来。

（3）在同一架唱机上，用别的唱片就会播出别的声音来。

（4）唱片和唱机很容易彼此分开。

DNA复制生物体的过程依赖极为复杂的细胞内的化学过程。这些过程并没有录制在DNA上。因此，可以有两种截然不同的观点。一种观点认为，生物复制的大部分信息是在DNA之外的机制中，DNA只是起了一种开关的引发作用。另外一种观点则认为，所有的遗传信息都贮存在DNA内部，只是采用一种极复杂的方式。按照第一种观点，要揭示DNA的含义离不开相应的生化过程。按照第二种观点，只要有足够的智慧就可以揭示DNA的内在含义。

一般来讲，揭示一个形式系统内部结构的含义就是寻找受形式规则控制的符号与现实世界之间的同构。对于比较复杂的同构，需要较多的“设备”才能从形式符号中抽象出意义来。这些设备包括硬件和软件。如果这种同构非常简单，或者很熟悉，那么我们就会因为它的显而易见或者习以为常而只看到符号的意义，却忘记了产生这种意义的同构。最明显的例子就是人们的日常语言。

我们现在设想，有一张巴赫奏鸣曲的唱片被带

出了太阳系，甚至被带出了银河系。如果那儿有高等智力的生物得到了它，一定会被它的形状吸引住，竭力想揭开它的秘密。即使它们的智力暂时还做不到这一点，他们也会在原则上相信，这里一定包含着某种信息。这种信念就是建立在破译的基础上。

今天，关于破译的思想在天文学、语言学和军事等领域已经被广泛地接受了。有一个突出的例子，就是破译用人们不懂的语言和字母写成的古文。我们本能地感觉到，在这种古文中一定包含着某种信息。而且，一旦这种古文的语言被破译了，没有人会怀疑，这些符号的意义究竟隐藏在哪儿。显然这些意义是在原文内，而不是在破译的方法之中。这就象乐曲是在唱片里而不是在唱机内一样。

我们可以将一分包含意义的形式符号称为消息，并将这分消息分为3个层次：（1）结构消息；（2）外部消息；（3）内部消息。

结构消息是通过载体的总体结构含蓄地传递的。它告诉人们："我是一份消息，如果你有能耐就来破译我。"理解了结构消息就是承认有破译的必要性。

认识到破译的必要性就转入第2个层次，即外

部消息。这种消息包含在消息的符号形式和结构里，它能告诉我们怎样理解内部消息。理解外部消息就是构造或者了解翻译内部消息的方法。

内部消息是我们最终的目标，是这份消息所要传递的信息。例如我们在音乐中体验到的激情、基因所产生的表现型特征、碑文中记载的古代文明中的王权或仪式等等。理解内部消息就是要把传递的意义提取出来。

为什么其他星球上的高等生物看到唱片就会想到其中一定隐藏着什么信息呢？首先就是因为唱片的几何形状，其次是因为唱片上螺旋状的非周期性纹沟。如果把这种盘旋的纹沟展直，就象是一条一千多米长的手写体符号。DNA 的分子形状也与此相仿。不过它用的“字母”是 4 种不同的碱基。其实早在艾弗里确定DNA与遗传信息的关系之前，物理学家薛定谔就已经预见到遗传信息贮存在“非周期性的晶体结构”中。这就启示我们，如果在非常规则的几何构形中发现了非周期性的结构，那就很有可能隐藏着某种内部消息。图24为此提供了一些出色的例子。

这 3 种层次在海边捡到的瓶子上也可以明显地表现出来。当人们捡起瓶子，看见它封着口，里面

又装有干燥的纸条，马上就会了解消息的第一个层次。然后打开瓶盖，取出纸条，查看上面所写的符号。也许这是用日文写的。这时并不需要知道它的内部消息，只要从字母形状的某些特征就可以识别出来。最后一步才是用日语去理解纸上所写的内容。

无线电短波的监听者也会面临同样的问题。首先他要确定自己听到的声音是一种消息而不是静电干扰。然后判定它的语种或电码。最后才是了解它的内容。

这些例子似乎证明了这样的论点，没有一种消息本身就有意义。哪怕最简单的消息，要想了解它的内部消息，首先就必须了解它的结构消息和外部消息。这就是说需要借助于“自动电唱机”，要把唱机的一部分信息加到这种消息上去才能理解它的意义。

这看来好象是一种恶性循环。你要想理解任何一份消息就必须先有一种消息告诉你如何去理解它。但是，我们都知道，这不是一种无意义的循环。因为我们确实能够理解消息。这是怎么回事呢？

因为我们的智能并不是抽象的东西，而是具有具体的物质基础：大脑。大脑的结构是长期进化过程

的产物，它的功能受到物质运动规律的支配，它们是物质性的实体。因此，我们的大脑在运行时并不需要有谁来告诉它们应该怎样运行。看来，大脑配备着“硬件”，可以识别哪些东西是消息，而且能够破译这种消息。这种最起码的提取内部消息的能力使得掌握语言的过程能够象滚雪球一样进行。因此，这种先天性的硬件就象唱机一样，为我们理解消息提供了辅助的消息。

由于大脑具有相同的结构，这种人类的自动电唱机也是类似的。这就为统一的“语言”提供了基础，从而使得结构消息和外部消息可以在人们之间进行交流。

最后，让我们来看这样一个例子。假设我们看到有一块方整的金属板上刻着上下排列的两个点。（见图版25）虽然我们猜测这里可能隐含着什么信息，但是很难发现它。如果我们看到一块金属板上刻着这些点：那就很容易发现，这就是费波那奇数列。实际上，从这样一些数可以有把握地推出费波那奇数的定义。如果我们把最初的一对值（1，1）看成是“基因”，那么按照费波那奇数的递归定义就可以展开整个数列（表现型）。但是仅仅给出“基因”（第一块金属板），我们无法获得重构表现型所需

要的信息，即级数的定义。而从更长的“基因”（第二块金属板）运用智能就可以得出由基因展开成表现型的机制。一旦获得了这种机制，那么从短的基因也能展开成表现型的级数。

2.4 音乐和绘画中的同构

同构不仅赋与形式符号以生命，也是音乐和绘画中的灵魂。巴赫在《音乐的奉献》中得心应手地驾驭的卡农技巧就是一种同构。即导句与伴句之间的同构。

卡农的花样繁多。当伴句与导句的音程完全相同或者相差 8 度时，称为同度卡农。如果伴句比导句低 5 度音程称为 5 度卡农。如果伴句高于导句 4 度音程称为 4 度卡农。伴句与导句可以有时间上的差异，也可以通过速度的变化来分出层次。速度加快为增时卡农，反之则为减时卡农。

有一种卡农结构是在伴句中把导句的主题变为它的倒影。这看来似乎是一种古怪的旋律交换，然而在听众的耳朵里却显得十分自然。还有一种巧妙的卡农结构是在伴句中把导句的主题按照时间顺序颠倒一下。人们把这种卡农亲昵地称作蟹式卡农，

即逆行卡农。（图27）尽管卡农的形式多种多样，但是万变不离其宗。在每一种形式的卡农中，伴句都保留了导句主题的全部信息。从这个意义上讲，由每一个伴句都可以使导句的主题完全复原。

充分运用卡农技巧的《音乐的奉献》是巴赫的杰作，也是一个值得人们不断地去发掘的宝库。其中有一支三部赋格曲、一支六部赋格曲、十部卡农和三重奏鸣曲。这支十部卡农是巴赫写过的最复杂的卡农。奇怪的是巴赫并没有把它全部写出来。这就好象是给弗雷德里希国王出了一个意味隽永的谜。

埃舍尔则把卡农变成了图画。（图13）看到这幅妙趣横生的《“逆行卡农”》，读者也许会联想到各种图案设计中形形色色的“卡农”。图形的平移是有音程差的“卡农”。图形的镜象变换则是倒影“卡农”。用数学的语言来描述，这些图案是通过对称性的交换而构成的。这些变换的共同特点是同构性。这些变换的乘积仍然是一种同构交换。这也表明同构的关系是可以传递的。

图案变换的同构性还表现在它们的“相似性”上。这里说的“相似性”并不是指几何学中图形的相似性。让我们看看《蝴蝶》（图12）中的“相似

性”。把一只蝴蝶映入另一只蝴蝶并不是严格地保持各部分之间的比例关系，而只是保持某种相互关系。更确切地讲是保留某种相互关系的信息。而这就是同构。

3

形 式 系 统

3.1 从一个数学难题谈起

为了进一步阐述哥德尔定理的内容，我们必须和一些抽象符号构成的形式系统打交道。也许有的读者对于这些抽象的表示方法感到不习惯，甚至不耐烦。不过我们还是希望读者有耐心去慢慢熟悉这种抽象的形式，这对我们在深一层次上理解哥德尔理论是必不可少的。我们所面临的是一座形式符号的大厦。形式符号是这座大厦的砖瓦，它们代表着公理和定理。形式推理的规则是这座大厦的框架，体

现了整个系统的结构。我们把整座大厦称为形式系统。其实在上章阐述同构概念时，我们已经简单介绍了形式系统的概念，并且给出了一个具体的形式系统，即pq系统。为了使读者对于形式系统有更切实的理解，并引起一些读者对它的兴趣，我们在这里提出一个有关形式系统的数学难题。有兴趣的读者可以认真地思索一番，甚至拿出纸笔演算一番。

我们先构造一个MIU系统。这个系统中只有3种符号，即M、I、U。这些符号构成的串称为系统中的定理。先给定一个定理为MI，也可以说这是公理。现在要问："你能否根据下面的4条规则由MI产生MU？"这4条规则是：

规则1：如果一个串的最后一个符号为I，则可以再加上一个U。

规则2：如果有一个串为Mx那么可以再加上x而生成Mxx。这里的x代表任何一个由M、I、U组成的串。

规则3：如果串中出现连续的3个I，那么可以用U代替III而得到一个新串。不过不能用III去代替U。

规则4：如果串中出现UU，那么可以把UU删去。

我们应该如何来考虑这个难题呢？我们可以设

想有一个长生不老的妖怪，他热衷于在MIU系统中，运用上面几条规则来生成各种串。

步骤1：把4条规则应用于公理MI，由此生成2条新串：MIU，MII（规则3，4不适用）。

步骤2：把4条规则应用于步骤1生成的新串MIU、MII。由此产生了3条新串MIIU、MIUIU、MIIII。

……

采用这种方法迟早会产生系统中的每一条定理。（见图版27）而我们要解决的难题是，MU是不是该系统中的一条定理？一般来讲，这就是检验定理的问题，或者说定理的判定问题。用构成系统中所有定理的方法来判定某一条串是否为系统中的定理，这并不能保证，这个过程以有限的步骤来完成。而这恰恰就是问题的关键所在。

在我们所出的难题中，只有一条公理，推导的规则是简单明了的，构造定理树也不困难。按理说定理的判定是可以做到的，但是判定的结果并不是显而易见的。不少读者大概仍然无法判定，MU是不是系统中的一条定理。

不过问题的解决往往取决于你从什么角度去看，也就是找到合适的同构机制。如果我们把注意

力集中在I的数目上，那么就会发现，这个MU难题实质上是一个用语言掩盖着的自然数难题。如果我们计算每个定理中I的数目就会发现，它好象永远不会为0。不但如此，我们可以进一步证明，I的数目永远不是3的整数倍。

我们先来看规则1和4。它们保持I的数目不变。然后看规则3，它永远不会凭空使I的数目为3的倍数。这就是说，在原来的定理中，如果I的数目不是3的倍数，那么经过这种变换后仍然不是3的倍数。只有在作为输入的定理中，I的数目为3的倍数，才能保持这种性质。规则2也是这样。要使加倍后的数能被3整除，就需要原来的数也能被3整除。总之，这些规则都不能使I数目不为3的倍数的定理凭空产生这种性质。

公理中I的数目为1。按照刚才的分析，运用这些规则决不会使新定理中I的数目为3的倍数，当然也就不能为0。因此MU不是MIU系统中的定理。

当然，并不是所有的这类问题都能这样容易地解决的。但是我们至少已经看到了，有一个难题可以归结于数论中的问题而得到解决。我们还将进一步看到，有一种方法可以将所有形式系统中的问题

归结于数论中的问题。这要归功于哥德尔所创造的一种特殊的同构，即哥德尔编码。

我们以MIU系统为例，可以在系统的符号与数字之间建立这样的对应关系：

$$M \Longleftrightarrow 3$$

$$I \Longleftrightarrow 1$$

$$U \Longleftrightarrow 0$$

这种对应关系完全是任意的。我们称这些数为哥德尔数。

采用哥德尔数，就可以从两种不同的层次去理解系统中由数字构成的串。一方面，可以把它们的运算看成是定理的变换；另一方面，又可以把它们看成是一般数字的运算。

现在设m、n为任意的自然数。于是这几条规则的算术运算就可以表述成：

1.如果生成了$10m+1$，就可以生成$10\times(10m+1)$。

2.如果生成了$3\times10^m+n$，就可以生成$10^n\times(3\times10^m+n)+n$。

3.如果生成了$k\times10^{m+3}+111\times10^m+n$，就可以生成$k\times10^{m+1}+n$。

4.如果生成了$k\times10^{m+2}+n$，就可以生成$k\times$

$10^m + n$。

公理则可表述为：　　能生成31。

这样的算术系统称为310系统。在这个算术系统中，可以生成的数起着形式系统的定理的作用。因此，问题就归结为，在这样的系统中能否出现30这个数？按照我们已经作出的分析，这是不可能的。

如果你对这个问题进行了认真的思索，或者动笔进行了演算，那么你一定对于形式系统有了感性的认识。你懂得了形式符号、定理和形式系统三者是什么关系。什么是形式系统的规则和公理，它们如何决定了在系统中可以生成哪些定理。你也知道了公理和定理的区别仅仅在于，公理是任意规定的，而定理则是由公理按照规则生成的。所谓形式推理是严格按照形式系统的规则，逐步生成新的定理，直到我们所需要的定理出现为止。这个过程也叫做证明。其实，我们对于形式系统并不陌生。我们可以回忆一下中学时代的几何学。所谓几何学命题的证明就是根据已知的几何定理和逻辑法则（起形式系统规则的作用），从条件（已知的定理）推出所要证明的结果（需要证明的定理）。人们把这种训练称为“思维的体操”。看起来单调的体操动作可以使我们的体格强壮。看起来乏味的形式推理对于我

们的思维也是不可缺少的。希望读者也能逐渐熟悉和运用形式系统。

3.2 一致性和完备性

形式系统的推理是严格按照规则，从已有的定理产生新的定理。这种机械化的步骤可以使系统中的定理不断增加而形成庞然大物。形式系统是一座由公理、定理通过规则结合成统一整体的蔚为壮观的大厦。要使这座大厦能在狂风暴雨中巍然挺立，也能承受地震的考验，就必须满足一致性的要求。在这些定理中不能有相互矛盾、相互冲突的结果。否则这座大厦就会产生缝隙，从而使整座大厦倒塌。用逻辑学的语言来讲，一致性就是不容许在同一个系统中既出现一个判断为真的命题，又出现同一判断为假的命题。

当然，可以有各种各样的一致性。逻辑上的一致性要求避免逻辑上的矛盾，这是形式逻辑的基础。数学上的一致性不容许出现违背数学理论的那些描述。物理学上的一致性则要求对物理现象的解释与物理定律密切配合。还有生物学上的一致性等等。不过一般来讲，人们感兴趣的往往是数学与逻

辑的一致性同物理学上一致性的差别。

逻辑与数学的一致性也就是形式系统内部的一致性，就是说对于系统内所有定理的解释不会产生矛盾。物理学上的一致性则是指形式系统与外部世界的一致性。这两种一致性在形式系统中是密切联系在一起的。

埃舍尔的作品《相对性》（图 8）表现了一致性与矛盾之间的微妙关系。它那栩栩如生的实体感诱使你去寻找一种有关此图的解释，使得图中各部分的解释不会产生矛盾。你也许会感到这些楼梯很好玩。因为在同样的楼梯上，人们行走的方向并不一致。我们可以把某个楼梯看作我们进行解释的基础，看成一个“确定性的小岛”，然后以此为基础来解释与它有联系的各部分。再根据已经获得解释的部分与其他部分的联系进一步推广我们的解释。这很象是在形式系统中运用规则进行定理的变换。对于一般的图，这个过程很容易进行到底，从而使全图获得一种令人满意的解释。可是在这幅图中就不是那么称心如意了。我们很快就会遇到麻烦，产生冲突。如果我们放弃所选择的“小岛”，而另外选择一个“确定性的小岛”还会遇到同样的麻烦。但是，我们又无法否认这些东西的意义。这些楼梯

只能是楼梯而不可能是鱼、船或者什么别的东西。

这幅画告诉我们，虽然一致性是一个可以明确定义的概念，一个形式系统要保持一致性似乎是理所当然的。但是对于一个具体的形式系统来讲，一致性并不是显而易见的，对一致性的判断有时是很困难的。

如果说一致性是形式系统成立的基础，是保证一个形式系统中每个定理都有意义的必要条件，那么完备性就是对有意义的那些定理的最大程度确认。一致性表明形式系统的这种性质："该系统中生成的每个定理都是真的。"完备性则指出："每个为真的定理都在该系统之内。"完备性也叫完全性，同样不是一种显而易见的性质。欧几里得的《几何原本》提供了欧氏几何的公理系统，是数学理论中影响最深远的范本。但是它并不是完备的。希尔伯特在20世纪才把这个不完备的公理系统变成了完备的公理系统。他的不朽之作《几何基础》是一部重要的数学著作。希尔伯特明确地提出了关于公理系统的一致性、完备性和独立性的概念，并且以这些概念为基础来构造严密的数学理论。他是数学形式主义学派当之无愧的领袖。

3.3 形式系统的结构

我们知道了什么是形式系统中的定理和规则，有了一致性和完备性的概念，就可以进一步讨论形式系统的整体结构，讨论系统中的系统。还是以我们比较熟悉的几何学理论系统为例。

欧几里得的《几何原本》系统地总结了当时已知的平面几何、立体几何的知识。它在二千多年中实际上成了数学的圣经。欧几里得是数学推理严密性的奠基者。《几何原本》从最简单的概念、定义出发，逐渐构造起几何学的庞大体系。在这个体系中，每个结果都只依赖于前面已经得到的结果。这很象是层层叠叠的巴比仑塔式的建筑物（图7）。不过这座建筑的框架是抽象的。因此，如果有一条定理的证明不正确，并不会有明显的裂缝，也不会因此而使整个大楼倒塌。事实上，在《几何原本》的证明过程中使用了许多日常的语言。这是一种不严格的、有时是难以捉摸的通信媒介，里面有许多隐藏的陷阱。不过欧几里得是位深刻的思想家和数学家，你无法在他的著作中找到“跳跃”的漏洞。他以五条公理为基础，构造了无穷层的几何学摩天

大楼的最下面几百层楼。而且，他只用前面的四条公理推出了最初的28个命题。我们可以把这些命题说成是“四公理几何学”。人们常常把这称为“绝对几何学”。欧几里得显然有这样的愿望，能够用前面的四个公理推出第五公理，而不是把它作为一条公理直接加以规定。不过他没有成功，因而只好作罢。在他之后的许多世纪中，无数的数学家作过同样英雄的尝试，也都失败了。

但是后来经过萨谢利、兰伯特、波约、罗巴切夫斯基和高斯的努力，终于发展了几种形式的非欧几何学。他们并不是直接否定第五公理，而是抛弃这个公理，即平行公理的假设，而代之以其他形式的假设。欧几里得平行公理假设，在平面上过一点只能引一条与已知直线平行的直线。非欧几何或者假设过一点可以引无数条具有这种性质的平行线，或者假设过一点无法引任何一条具有这种性质的平行线。而按照这两种不同的假设，都可以构造出无逻辑矛盾的几何学体系。对于新的几何学体系中的各个命题，同样可以作出合理的新解释。

在形式系统中，有些符号是加以定义的，有的却没有。我们已经提到，在欧几里得几何学中，有许多未加定义的术语。因此我们可以把欧氏几何与

非欧几何中的术语分为两类。一类词汇具有固定的定义，是不能变通的。另一类则是未加定义的术语，它们的意义可以加以调节，从而保持整个系统的一致性。几何学需要有些第一类的词汇，它们是在几何学之外加以定义的。这些词汇构成了框架，而在这个框架中可以填充其他的材料。既可以填进欧几里得几何的材料，也可以填进非欧几何的材料。

因此，我们可以把某些形式系统分成同一系列的不同等级。一个形式系统可以把另一个形式系统包括在内，就象欧氏几何中包含着绝对几何学一样。用数学术语来讲，这就叫做把一个形式系统嵌入另一个更大的形式系统。例如，我们把形式系统 1 嵌入形式系统 2 。那就意味着，形式系统 1 中的公理和规则在形式系统2中仍然成立，是构成新系统的基础。但是在形式系统2中可以增添形式系统 1 中所没有的新公理或新的规则。需要补充的一点是，在一个形式系统中增加新的定理未必能构成能力更大的系统。关于这一点以后还要进一步加以阐述。

3.4 命题演算系统

我们已经了解了形式系统的一般结构，也构造了一些具体的形式系统，如pq系统、MIU 系统等。现在我们要介绍一种重要的形式系统，即命题演算的形式系统。这是逻辑命题形式化的结果，也是形式逻辑的重要组成部分。读者可以在任何一本有关形式逻辑的教材中找到有关这个系统的详细介绍。由于各位逻辑学家的偏爱不同，所用的符号可能有所不同。但是它们的含义并不因此而受到影响。运用我们前面讲到的同构方法，读者很容易在这些不同的符号系统之间建立一一对应的关系。我们这里采用的符号表是：

〈 〉

P Q R ′

∧ ∨ ⊂ ～

［ ］

我们把这里的P、Q、R称为原子，可以在原子的右上角加撇“ ”而构成新的原子。例如R′、Q″、P‴′。我们定义这些原子是系统中良构的命题。

可以按照下面 4 种方法由良构的命题生成新的

良构命题。

（1）～x　　x的否定

（2）〈x∧y〉　　x与y的合取

（3）〈x∨y〉　　x与y的析取

（4）〈x⊃y〉　　x蕴含y

这里的x、y都表示良构的命题。任何良构的命题都可以按照以上4种定则回归到最基本的要素——原子。从这种方法我们可以判定任何一个命题是不是良构的。这种过程一般称作是自上而下的。

这里的“〈”和“〉”是一组括号，表示括号中的逻辑符号构成了一个逻辑命题；“～”表示逻辑否定；“∧”表示逻辑积或者逻辑合取；“∨”表示逻辑和或逻辑析取。“⊃”表示逻辑蕴涵。逻辑中的否定、合取、析取和蕴涵可以根据原命题与新命题的真假值表来定义，这是形式逻辑教材的基本内容。

对于原子我们无法单独作进一步的解释，可以把它理解为任何一个用日常语言描述的句子。但是命题演算并不关心这些句子的实际含义是否正确。例如“这支笔是红的”就是一个原子。这支笔实际上可能是绿的。但是在命题演算中，这一点无关宏旨。重要的是以确定的符号来表示这个命题，并对

这些符号进行形式运算。这种运算是严格的、机械的，也可以说是无聊而乏味的。

由于命题演算系统的这种特点，因此在这个系统中只有运算的规则，而没有公理。按照前面几个形式系统的性质，我们知道如何按照规则从公理产生定理，再由这些定理产生新的定理。但是在这个系统中却没有公理。那么定理又从何而来呢？这很象是一个让巧妇做无米之炊的难题。幸好在这个系统中有一个定则可以使我们能“无中生有”地构造出系统的定理来。我们把这个定则称为想象定则。运用想象定则时，你可以先写下一条良构的命题，不妨设这个命题为x。然后设想x就是公理或者定理，再运用系统中的规则对它进行变换看看能得到什么样的结果。假设最后得到的结果为y。那么退出想象后就可以得到这样的定理：如果命题x成立，那么命题y也成立。用逻辑符号来表示就是〈x⊃y〉。

我们用方括号的前后两部分分别表示“进入”或“退出”想象。例如：

[	进入想象
<P∧Q>	前提
P	分离律

Q	分离律
<Q∧P>	结合律
]	退出想象
<<P∧Q>⊂<Q∧P>>	定理

这里只有最后一行是命题演算系统中的定理，其他仅仅是一种设想。这种想象定则可以重复地使用，从而使定理中的命题分出不同的层次来。我们不妨把这些层次称为“现实性的层次”。运用这个概念可以描述命题演算系统中的贮存定则：在想象中可以运用“现实性”更高的层次中的任何定理。

命题演算系统中的其他定则可以归结如下：

结合律：如果x和y成立，那么<x∧y>也成立。

分离律：如果<x∧y>成立，那么x和y同时成立。

双重否定律：“～～”可以在定理中删去或者添上。

独立律：如果x和<x⊃y>成立，则y也成立。

对照律：<x⊃y>与<～y⊃～x>可以互换。

德·摩根律：<～x∧～y>与～<x∨y>可以互换。

斯威彻罗律：<x～y>与<～x⊃y>可以互换。

要想验证这些定则，只要分别列出各个命题的真假值表就可以了。根据这些定则，可以产生命题演算系统中的各条定理。这些定理在有些人看来好象是纯思辨的，没有任何实际内容。按照这种观点，也可能把命题演算看成是白费时间的运算。因为它告诉我们的只是一些没有实际意义的东西。但是也有人会认为，它描述了所有命题的运算形式，它揭示的是对各种命题都普遍适用的规律，是更为深刻的真理。

形式系统要求严格按照规则来推导定理，或者说证明一个定理。但是怎么能够证明这种证明是正确的呢？这有点象一个鸡蛋。虽然有蛋壳的保护，但是在运输时，总要设法再加保护，使蛋壳不致于破碎。然而，不管保护层有多少层，总还存在这样的危险，有某种意想不到的灾难会使蛋壳破裂。当然这并不意味着，你永远不要去冒运输鸡蛋的风险。同样，你无法对形式系统证明的合理性给出一种绝对的证明。你可以给出一种证明的证明，一种证明的证明的证明。但是，仍然存在着未加证明的假设，我们人为地把它看成是真理。

例如，我们在命题演算系统中推出了$<P\wedge\sim P>$，现在要推出$<Q\wedge\sim Q>$。我们知道，这两

个推导过程是平行的，那就可以直接得出来。这在平面几何的证明中是经常碰到的。“同理可证”对于中学生是不陌生的。我们不妨把这称为元定理，是关于系统中定理的定理。运用元定理可以使形式系统繁复的推理过程大大简化，不失为一种人们乐于采取的捷径。

我们也可以把元定理看成是一个更大的形式系统中的定理。但是，人们仍然可以找到关于元定理的定理——元元定理。这就是更高层次上的推理捷径。

有时，形式推理的结论会产生明显的矛盾。这时我们往往会努力去弥补这些漏洞，使该系统更加完善。因此我们可以说矛盾是使理论系统达到清晰和获得进展的主要源泉。数学史上的著名例子就是有关无穷级数：

$$1 \quad -1 \quad +1 \quad -1 \quad +1 \quad -\cdots\cdots$$

之和的争论。人们“证明”它可以等于0，1，$\frac{1}{2}$或许还有其他值。这显然是很荒唐的。这种争论后来导致关于无穷级数收敛性的认识，从而发展了关于无穷级数更加深刻、更加全面的理论。

命题演算是模仿我们的逻辑思维，它与我们的

实际思维方式之间存在着很大的差异。这使许多逻辑学家感到不安，他们作出创造性的努力去补救这一点。这种努力推动了逻辑学的发展，使命题演算具有更大的灵活性。也有的逻辑学家甚至采取更加极端的立场，为了模仿人类的实际推理方式，宁可放弃对于完备性或者一致性的要求。

3.5 形式数论系统

最后，我们要讨论一种本书中最复杂但也是最重要的形式系统，即形式数论系统，简称 TNT系统。它也是罗素和怀特海在《数学原理》中提出的形式系统。他们提出这个系统的目的是要用它囊括现有的全部数学理论。当然，这个目标后来被证明是永远达不到的。

形式数论使数论的所有命题形式化。为此我们先来看一些典型的数论命题：

（1） 5是素数。

（2） 2不是平方数。

（3） 1729为两个立方数之和。

（4） 没有两个正立方数之和是另一个立方数。

（5）存在无限多的素数。

（6）6是偶数。

我们将这些陈述分解成为一些基本的术语：

对于所有的数b

存在着数b使得

大于

乘以

加上

0，1，2……

当然，这种分解并不是唯一的，例如“大于”可以进一步分解。a大于b可以表示成：

存在着数c不为0，使得c加b等于a。

分解成基本术语后，就可以进一步使这些术语形式化了。我们用这样一些符号来表示自然数。

数：	0	1	2	3	……
符号：	0	S0	SS0	SSS0	……

这里的S表示后继者，这就是说任何自然数可以用0和后继者的概念来描述。也许有人对于这种形式的表示方法感到不习惯或者不耐烦。不过我们在前面已经讲过，形式化为获得严密性而付出的代价是单调、乏味和冗长。不过，如果人们想到，这将更加便于计算机来“理解”，也许会理解形式化的必要

性。

和代数中一样，我们用字母a、b、c等来表示变量，用右上角加撇的方法则可以使变量的数目无限地增加。

加、乘、等于都可以用大家熟悉的符号“+”“·”和“=”来表示。此外还要加上两个重要的量词符号：

∃b：表示存在b使得　　（存在量词）

∀b：表示对于所有的b都有（全称量词）

出现在可真可假命题中的变量，我们称为自由变量。就是说它的某些取值使命题为真，而另外一些取值使命题为假。受量词支配的变量则称为量词限定的变量，上面的b就是其中一例。

有了这些术语和符号，我们就可以使数论中的那些陈述形式化了。就以我们举出的那些数论陈述为例：

（6）6是偶数。

$\exists e$：　$(SSo \cdot e) = SSSSSS_0$

（直译为：存在着e使得2乘以e等于6）

或者$\exists e$：　$(e \cdot SS_0) = SSSSSS_0$

或者$\exists e$：$SSSSSS_0 = (SS_0 \cdot e)$

（2）2不是平方数。

$\sim\exists b:\ (b\cdot b) = SS_0$

（直译为：不存在b使得b乘以自己等于2）

或者$\forall b:\sim (b\cdot b) = SS_0$

（直译为：对于所有的b都不会自己相乘等于2）

（3）1729是两个立方数之和。

$$\exists b:\ \exists c:\ \underbrace{SS\cdots\cdots SS}_{1729个S}{}_0$$

$$= (((b\cdot b)\cdot b) + ((c\cdot c)\cdot c))$$

（4）没有两个正立方数之和为另一个立方数。

$$\forall a:\ \sim\exists b:\ \exists c:\ ((a\cdot a)\cdot a)$$

$$= (((Sb\cdot Sb)\cdot Sb) + (Sc\cdot Sc)\cdot Sc))$$

或者$\sim\exists a:\ \exists b:\ \exists c:\ ((a\cdot a)\cdot a)$

$$= (((Sb\cdot Sb)\cdot Sb) + (Sc\cdot Sc)\cdot Sc)$$

这里之所以用Sb、Sc而不用b、c，是为了保证它们为正整数。

（1）5是素数。

$$\sim\exists b:\ \exists c:\ SSSSS_0 = (SSb\cdot SSc)$$

这里之所以用SSb、SSc是为了保证它们大于1。

（5）素数是无限的。

$$\forall d:\ \exists e:\ \sim\exists b:\ \exists c:\ (d + Se)$$

$$= (SSb \cdot SSc)$$

（直译为：对于任意数d存在着数e使得d + e + 1不是任何两个数（b + 1）与（c + 1）的乘积。）

从这些例子我们可以清楚地看到，运用上面例举的那些符号，就可以使得数论中的那些陈述形式化。

皮亚诺建立了关于自然数的公理体系。我们可以把他的公理体系归结为3个基本概念和5条公理。这3个基本概念是0、数和后继。5条公理是：

（1）0是一个数。

（2）任何数的后继是一个数。

（3）没有两个数有相同的后继。

（4）0不是任何数的后继。

（5）任何性质，如果0有此性质；又如果任一数有此性质，它的后继必定也有些性质；那么所有的数也有此性质。

皮亚诺的公理系统对于数学理论的形式化具有很大的影响。我们可以在这里清楚地看到它对形式数论系统的影响。

我们还可以把命题演算系统中的所有规则都引进TNT系统来。也就是说用它们来构成TNT形式系统的框架。这就意味着，在形式数论系统中的推

理必须遵循形式逻辑的法则。

根据TNT的规则和公理，我们很容易生成这样的塔状定理族

$$(0+0) = 0$$
$$(0+S0) = S0$$
$$(0+SS0) = SS0$$
$$(0+SSS0) = SSS0$$
……

把这族定理归结在一起，就可以得出采用全称量词的公式：

$$\forall a: \quad (0+a) = a$$

我们用这个例子来给出关于ω－不完备性的严格定义：

如果在一个系统中，某个塔状族中每一个定理单独都成立，可是用全称量词概括的定理却不成立，那么这个系统就是ω不完备的。如果这个全称量词的公式的否定也不是该系统中的定理。那就是说，原来的公式在该系统内是不可判定的。这看来好象不可思议，其实并没有什么神秘之处。这只是表明一个形式系统是可以扩张的。欧几里得第5公理即平行公理在绝对几何学（4公理系统）中是不可判定的。但是加上这个公理后便成了欧氏几何。

而加上相反的命题后便成了非欧几何。

但是，如果在这种塔状族的定理中，每一行都可以从它的上面一行按一定的规则推导出来。那么我们完全有理由相信，这族中的所有定理都成立。

我们用X{a}表示带有自由变量a的良构的公式。X{Sa/a}则表示用Sa代替原公式中的a。那么上述情况就可以用TNT中的一条定则来表示。我们把这条定则称为归纳定则。

归纳定则：若u是自由变量，X{u}是以u为自由变量的良构公式。如果∀u：〈X{u}⊃X{Su/u}〉和X{o/u}都是定理，那么∀u：X{u}也是定理。

显然这是皮亚诺第五公理在形式数论系统中的表述。在TNT系统中还有这样一些公理。

(1) ∀a：～Sa = 0

(2) ∀a：(a + 0) = a

(3) ∀a：∀b：(a + Sb) = S(a + b)

(4) ∀a：(a · 0) = 0

(5) ∀a：∀b：(a · Sb) = ((a · b) + a)

这些公理的含义是很容易理解的。公理1陈述了有关0这个数的专门特点。公理2和3是有关加法运算的。公理4和5是有关乘法运算的性质，而且还

涉及到与加法运算的关系。

我们不想在这儿详细叙述形式数论系统的所有公理和规则。这将会使许多读者感到厌烦。希望读者通过这些说明获得一个明确的印象，有关数论的命题和规则能够通过同构而形式化。

4

递　　归

4.1 阿基里斯和乌龟历险记

我们在上一章介绍了形式系统，并在最后介绍了形式数论系统，即TNT系统。TNT系统的一条定则就是归纳定则，也就是递归的定则。递归是极为重要的概念，它对于理解哥德尔定理的证明是必不可少的。为了说明这个重要的概念，我们先从一段故事讲起。

（乌龟和阿基里斯在科尼岛上消磨时光，他们

坐在大转轮上谈笑风生。）

龟：我有一种预感，今天幸运之神将会赐福于我。

阿：这可太好了。你看！有一架直升飞机正朝着我们飞来了。

龟：它还放下了一段绳子。看！快到我们身边了。

阿：啊，绳子上有一个大钩子，上面挂着一张纸条。

（他们按照纸条上的吩咐，松开自己坐椅上的皮带，在轮椅再度上升到顶的一刹那抓住钩子上了直升飞机）

声音：哈、哈、哈！朋友们，你们上钩了。先在我的书房里呆着吧。不妨吃点爆米花。等我磨好刀子就来收拾你们。乌龟馅饼，这可是我最喜爱的食品。

（声音消失后，乌龟阿基里斯在书房里面面相觑。忽然他们发现了一本名叫《阿基里斯和乌龟历险记》的书。

龟：这可真是一个富有刺激性的题目。我们读一读怎么样？你扮演书中阿基里斯的角色，而我扮演书中乌龟的角色。

（他们开始读这本书）

〔在书中：

（阿基里斯邀请乌龟参观他所收藏的画家埃舍尔的作品。）

龟：这些作品真迷人！我最喜欢那幅《凸与凹》（图9），那里有两个内部一致的世界，可是结合在一起便成了不一致的复合世界。如果能访问一下这种不一致的世界倒是非常有趣的。不过，我可不愿长久地生活在这样的世界里。

阿：你要知道，不一致的世界是根本不可能存在的。你怎么能去访问这样的世界呢？再说，埃舍尔的画是一个二维的世界，你怎么能进得去呢？

龟：我自有办法。我只要喝一小杯神奇的“进入剂”就可以做到这一点。如果我想从进入的画面再退出来，只须喝一小杯“退出剂”就可以了。

（阿基里斯与乌龟倒了两杯“进入剂”并带上“退出剂”一起干杯。）

阿：我们这是到了哪儿？在一只小小的平底船上。船顺着运河而下。喂，船夫，快让我

们下船。

（可是船夫根本不理睬他们。他们终于明白了，船夫听不懂他们的话。他们来到了一个完全陌生的国度，就只好在这个希奇古怪的世界里乱闯乱窜。后来阿基里斯意外地得到了一盏铜灯。他慢慢擦着灯上镌刻的字样。突然冒起一股浓烟。阿拉伯神话中的奇景再现了。灯神愿意满足新主人的三个愿望。可是阿基里斯却想满足上百个愿望。灯神不忍心让阿基里斯失望，就从自己的长袍中取出一盏形状相仿的银灯。）

阿：这是什么？

灯神：这是我的元灯，它可以满足人们的元希望。

（灯神摩擦元灯，在烟雾中出现了元灯神。元灯神又从自己的长袍中取出金制的元元灯。而元元灯神可以满足元元希望……这个无限的过程通向上帝。上帝可以许诺无形的希望，包括任何希望的希望。但是这种希望却是没有保障的。）

阿：这使我产生了一种特殊的希望。我希望自己的希望是无法实现的。

（灯神实现了他的愿望。于是灯神、神灯连同那些神奇的法力统统都消失了。整个系统都崩溃了。阿基里斯发现他们又来到了埃舍尔的另一幅画《蜥蜴》（图10）之中。）

［在《蜥蜴》画中

龟：看，在这串蜥蜴旁边居然还有一瓶“退出剂”。我们真算是有福气。

阿：我想最好还是从埃舍尔的画退回到我们自己的房子里去。

龟：这桌子上还有两本书呢，书的题目恰好也是《阿基里斯和乌龟历险记》。

（阿基里斯一心想离开这儿，他怕失去那瓶“退出剂”。可是他在慌乱中把那瓶药水碰翻在地、滚下楼去了。于是他只好和乌龟一起读起那本书来了。）

［在书的故事中

（阿基里斯和乌龟置身于迷宫的层层高墙之中。他们怀着求生的急切心情寻找着迷宫的出路。忽然他们听到了巴赫的变调乐曲《和谐的小迷宫》。正当他们被深深吸引住的时候，却听到了在迷宫中等候他们的魔鬼的笑声。他们处于极度恐惧之中却发现了一碗

爆米花。于是两个人就大嚼特嚼起来。突然砰的一声巨响。）

〔回到《蜥蜴》画中

龟：多么有趣的故事。

阿：我只关心他们最后能否逃脱那只魔鬼的手掌。可怜的阿基里斯，他可不想首尸分离。

龟：不必担忧，只要他们有“退出剂”就万事大吉了。

阿：可是那些蜥蜴从画中出来进去并没有喝什么“进入剂”或“退出剂”啊。我们是否也能仿照它们从埃舍尔的这幅画中退出去啊？

龟：当然可以，你跟着我一起做。

（于是阿基里斯和乌龟就从《蜥蜴》中退出来了。可是阿基里斯发现，自己并没有回到原来看画的书房，而是来到了乌龟的书房中。乌龟当然十分满意，阿基里斯也想就此收场。这可算得上是一段奇妙而幸运的经历。）

4.2 形形色色的递归

在上面这一节令人眼花缭乱的历险记中，阿基里斯与乌龟从一个世界进入另一个世界。这暗示着

一种重要的形式系统的结构，即递归结构。

这段历险记是一个很复杂的递归例子。阿基里斯和乌龟出现在各个不同的层次上。有时他们从一个层次进入另一个层次，有时则在读自己在其中扮演角色的故事。下面显示了这段经历的递归结构：

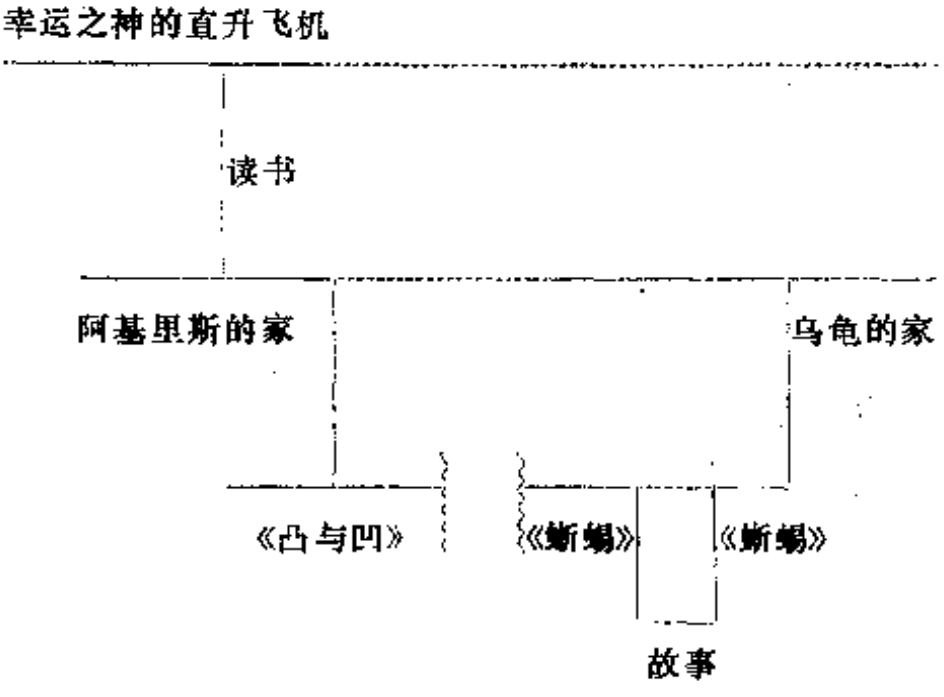


图30　“阿基里斯和乌龟历险记”的结构图

这张图显示了历险记中各个层次的结构。垂直地下降为“进入”，垂直地上升则为“退出”。从图中我们也可以看到，由幸运之神一开始的威胁所带来的紧张始终没有消退。阿基里斯和乌龟一直在直升飞机里晃悠。有些读者可能为此而感到不安，

也有人会无动于衷。尽管有这么多层次，但是它们都有一个共同的特点。在每一个层次里，主角都是阿基里斯和乌龟。这提醒我们注意，对于同样的符号可以从不同的层次去理解它的意义。

在我们生活中，递归的例子并不罕见。故事中的故事、画中之画、电影中的电影乃至中国式的套盒等等都具有递归的结构。在音乐中也可以找到递归的例子，象普柯菲耶夫的第5钢琴协奏曲和拉赫马尼诺夫的第2交响曲。

20世纪50年代，人工智能的语言就已经引进了“压进”、“退出”、“叠加式存贮器”这样一些相互关联的概念。“压进”意味着暂时停止目前进行的操作，但是并没有把它忘掉，而是去完成更低一层的任务。“退出”则正好相反，是结束在这个层次上的操作，回到更高的层次上来，重新开始因为“压进”而中断的操作。这些术语是受到食堂里放盘子的架子的启示。这种架子在盘子底下装配着弹簧，可以保持盘子最上面的高度不变。每当你压进一只盘子时，这叠盘子就下沉一层，而当你取出一只盘子时，它就上升一层。

递归过程也是电子计算机的基础。在计算机程序中，一种关键的技术就是在重复使用某一部分的

程序时可以采用模式化的方法。也就是说，把总任务分解成一些子任务。例如在进行一系列相同的操作时，就不必把它们全都写出来，而是画一个圈。这个圈就表示重复一组固定的操作，直到满足某种条件为止。因此，我们可以这样来理解圈，即重复地执行一系列预定的操作，而当预定的条件满足时，就中断这一操作。这实际上就是从一个层次进入另一个层次，然后又退回到原来的层次来。还有一种圈，我称它为“自由圈”。这种圈是很危险的陷阱。因为它要满足的条件可能永远不会出现。因此计算机一旦陷入这种圈里，就会进行无限的循环而无法解脱出来。把这种自由圈和有界圈区别开来在计算机科学中是很重要的。比圈更一般的概念是子程序。它的基本思想就是把一组操作结合在一起而形成一个单元。它们作为一个整体而被调用。

有一个典型的递归例子，这就是用计算机下棋时选择一种“最佳”的步骤。这就要对每一步以后的各种可能的步骤进行评价。要进行这种评价，就是设想在自己采取的每一步之后，对手有可能采取什么样的对策。而对于这样的每一种对策，自己又可以采取什么样的反对策。这样一步步地设想下去就是一种递归结构。一位高明的棋手与一般下棋爱

好者的区别之一，就是他能设想好几步棋之后的局势。这实质上是一种利用递归结构的优势。所谓“深谋远虑”也有这种含义。

我们在智力上的叠加存贮能力显然要比语言上的叠加存贮能力更强。在语言中也有一种压进－退出的叠加结构。但是这种结构的层次一多，理解句子的困难性就会大大增加。中国人往往喜欢较短的句子，这样文字就显得流畅。对于有些译文中出现的带有许多从句的长句子，有人则感到费解和不习惯。不过在英语中，使用从句增加句子的层次是司空见惯的，而在书面德语中，层次叠加的现象就更加明显了。即使这样，也总有办法重组这些句子使得结构的层次减到最少。

句子的结构为我们提供了一个很好的例子来说明递归的结构。这就是所谓的“语次转变”的递归网络。这种图的模式是由一些结点或者其中有单词的小方块以及带箭头的弧或者线组成。

图31(a)显示了一个语次转变的递归网络图，即带修饰的名词。可以有单独的名词，也可以有带形容词的名词或者是带冠词的名词，还可以是同时带冠词和形容词的名词。图31(b)则是一个想象名词的语次转变递归网络。在这个网络中我们可以看

到，每一条途径都必须经过“带修饰的名词”，而且有三条途径中都有“想象名词”。这看来好象是在用自己定义自己。但是，因为存在着从“带修饰的名词”到“结束”的途径，就保证了这个递归过程可以是有限的。

我们也可以这样理解有递归结构的网络图。这就是用整个“想象名词”的网络图来代替图31(b)中的“想象名词”这一格。我们把称为网络中结点的扩张。而扩张后的网络中的结点还可以再扩张。

几何结构的无穷扩张也可以采用这种逐层扩张结点的方式来定义。作为一个例子，我们来定义一种G图。在图32(a)给出了未经扩张的G图。图32(b)则给出了经过一次扩张的G图。图中从下往上依次标上了字号。这个无限的树具有一种十分有趣的性质。最右边的数字依次构成了费波那奇数列。这个数列最好是用一对公式来进行递归定义

$$FIBO(n) = FIBO(n-1) + FIBO(n-2), \quad n>2$$

$$FIBO(1) = FIBO(2) = 1$$

这种G图还有一种更令人惊奇的性质，它的结构完全可以用一个递归定义来表示：

$$G(n) = n - G(G(n-1)) \qquad n>2$$

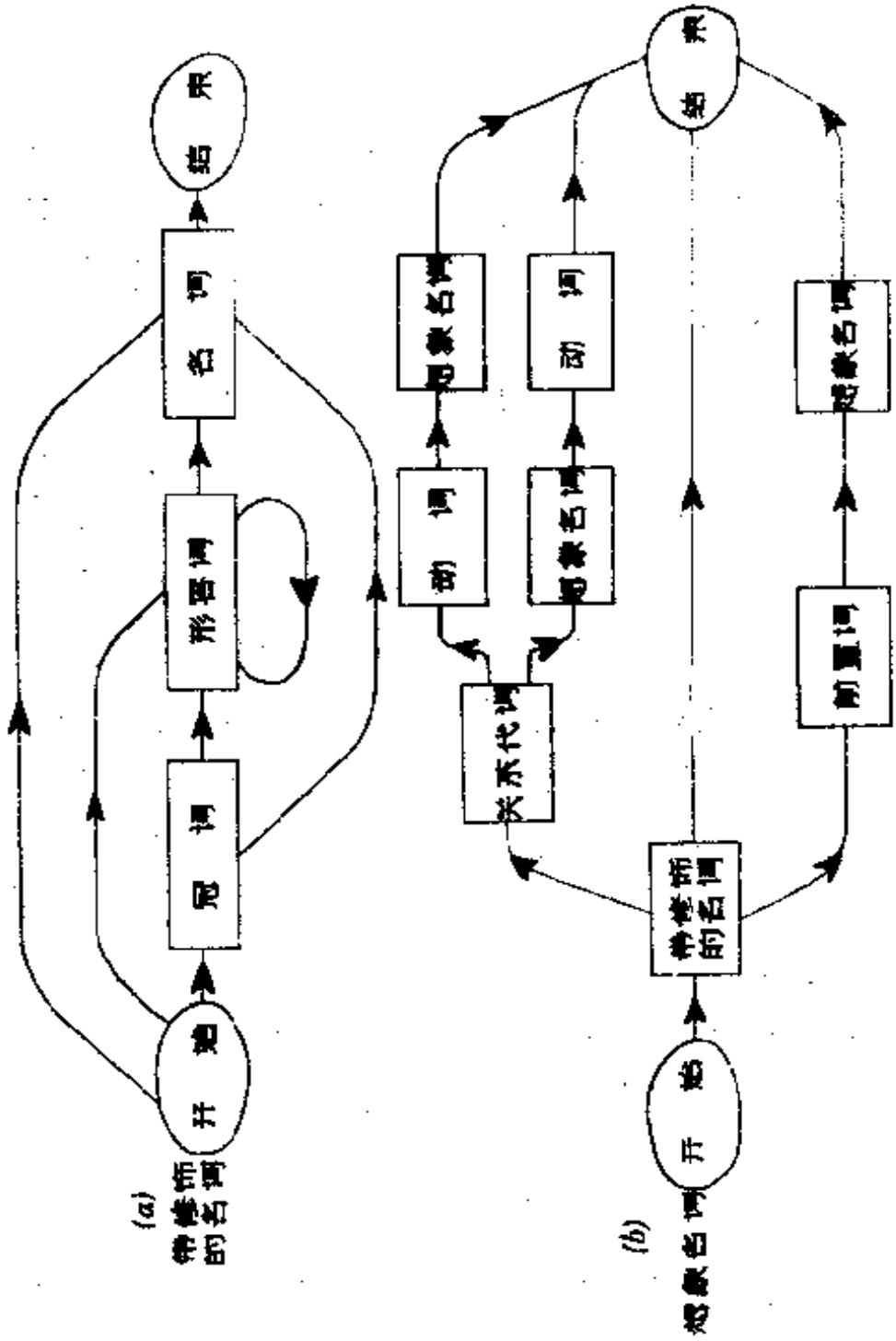


图31　语次转换递归网络：带修饰的名词和想象名词

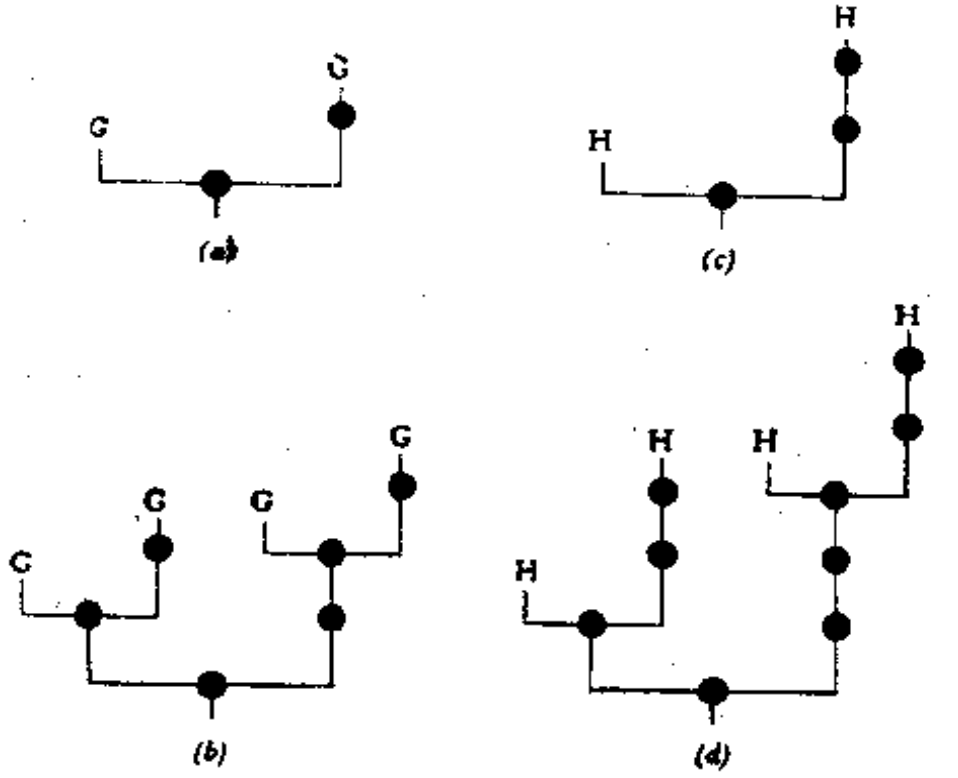


图32　(a) 未经扩张的G图
(b) 一次扩张的G图
(c) 未经扩张的H图。
(d) 一次扩张的H图。

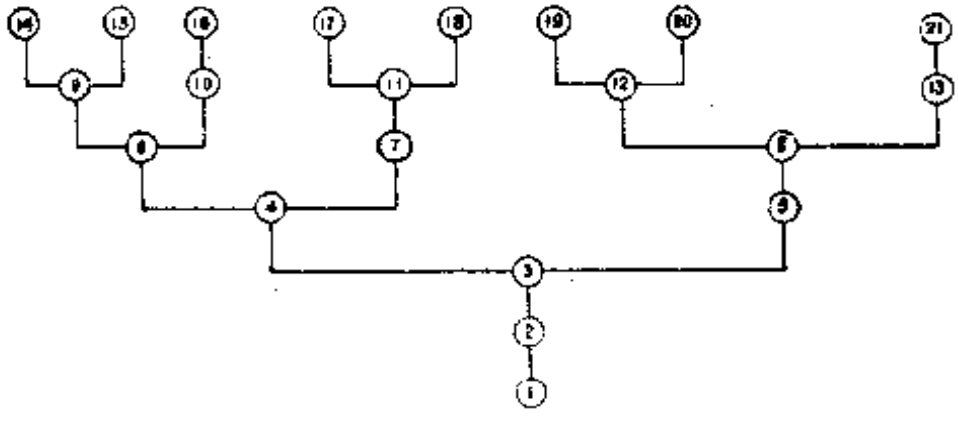


图33　进一步扩张并标有字号的G图。

$$G(0)=0$$

为什么函数G(n)能表示G图的树结构呢？非常简单，只要把G（n）放在每一个值n之下来构造一个树，就可以重新构造G图。实际上，我一开始就是用这种方法发现G图的。当我研究函数G时，为了迅速计算它的值，就把已经知道的值都标在树上，结果竟然得到了这样有序的几何结构。

同样，对于函数H(n)，可以给出这样的递归定义：

$$H(n)=n-H(H(Hn-1)))\qquad n>0$$
$$H(0)=0$$

有兴趣的读者可以思考这样一个问题：如果对G图作镜反射的变换，然后将树上的每个节点都重新标号，从左到右依次增加。那么，能否找到这种新图的代数递归定义呢？若将H图作这种变换，结果又会怎样呢？它们有没有共同的规律性？

还有一个有趣的问题是把两种递归定义“嫁接”在一起。例如：

$$\left.\begin{aligned}F(n)&=n-M(F(n-1))\\M(n)&=n-F(M(n-1))\end{aligned}\right\}\quad n>0$$

$$F(0)=1,\ M(0)=0$$

问题在于如何去发现F图与M图的递归结构。它们

是简单而优美的。

最后，让我们考虑这样一个递归的例子：

$$Q(n) = Q(n-Q(n-1)) + Q(n-Q(n-2)) \quad n \geq 2$$

$$Q(1) = Q(2) = 1$$

这使我们联想到费波那奇级数的定义，其中每一项都是前面两项的和。所不同的是，并非相邻前两项的和，而是按照这两项的值向左移动相应的位数，将由此得到的项的值相加得到新项的值。下面是该数列前17项的值：

1，1，2，3，3，4，5，5，6，6，

5＋6＝11

6，8，8，8，10，9，10，……

向左移动几位

第16、17位的值分别为 9 和10，这表示从第18位向左移动 9 位和10位，将由此得到的第 8 第 9 位上的值5和6相加就得到第18位的新值11。这样得到的数列看来是飘忽不定的，而且越是往后，越是难以发现它的规律性。这是一个颇为神秘的特殊例子。一种很自然的定义却导致了令人迷惑的行为：在一种非常有序的状态中产生了混乱。

在前面提到的费波那奇级数中，有两点是很重

要的。第一是递归公式，第二是初始值。如果我们改变一下初始值，使第一个值为 1 、第 2 个值为 3 ，那就会产生完全不同的数列，这就是罗卡斯级数：

1 、 3 、 4 、 7 、11、18、29、47、76、123……

我们谈到了语言中的递归结构、数学中的递归结构。现在再来看看一种图形的递归结构。在这种结构中，图形的整体和它的部分在结构上是相似的，虽然存在一定程度的变形和扭曲。埃舍尔运用这种思想创作了《鱼和鳞》（图11）。当然，这种鱼和鳞的相似性只有用一种抽象的方式来看才是相似的。谁都知道，一片鱼鳞并不是一条鱼的小付本，但是一条鱼的脱氧核糖核酸确实存在于鱼的每一个细胞之中，这是整条鱼的一个付本。这才是埃舍尔的《鱼和鳞》所揭示的深刻真理。

有时递归会给人一种错觉，好象是用自己在定义自己，从而导致无穷的回归。其实并不是这样。只要仔细作一番分析，就可以消除这种错觉。

4.3 图形和背景

自然数是人们最早认识的数学系统。关于自然数性质的理论——数论是最古老的数学分支之一。

许多杰出的数学家对于自然数或整数的许多奇妙性质特别感兴趣。说到自然数的一些性质，那就不仅仅指那些可以计算出结果来的性质，更主要的是指那些数学家们热心探索而又无法通过计算过程来加以证明的性质。

例如有一条古老的陈述："存在着无限多的素数"。这就是欧几里得定理。它并不是显而易见的。但是欧几里得之后的数学家都承认它是正确的。为什么呢？就是因为推理。依靠形式的推理加以证明就使我们不必去分别处理无穷多的实例。

在推理过程中，我们往往使用"所有"这样的词。这个词本身具有有限的形式，但却包含着一种无限。当我们运用形式系统中的规则时，总是认为它对"所有的"定理都是适用的。也许我们并没有意识到这一点，只是觉得我们是在按照这个词本身的意义来使用它。但是这实际上就意味着，我们是在被自己也不完全清楚的规则牵着鼻子走。在我们的一生中总是以某种特定的方式来运用词汇的。我们把这称为思维过程中词汇的意义。这一点对于探索数论的形式化具有至关重要的意义。

我们已经懂得如何通过同构来把握形式符号的意义，也可以通过形式符号的运算来把握概念。这

在有些人看来似乎不可思议。但我们确实在pq系统中以这种方式把握了加法的概念。而且这一切显得那样自然。

那么这种能力的限度究竟有多大呢？譬如自然数可以分为素数和合数。它们的意义是明确无误的。我们能否用形式系统的符号及其运算来判定某一个数是不是素数呢？

我们可以构造一个形式系统，其中的定理具有Px的形式。这里的x为连字符号“-”的串。只有当连字符号的数目为素数时才是该系统中的定理。

为了判定能否用形式符号的运算来达到我们的目的，我们先把前面几种形式系统出现过的运算归结一下，从而得到下面这张表：

（1）读出或识别有限的符号集合中任何一个元。

（2）写下属于这个集合的任何一个元。

（3）把任何一个符号从某一处复制到另外一处。

（4）消去这些符号。

（5）检验两处的符号是否相同。

（6）保存或使用前面生成的定理表。

这张表似乎显得有点累赘，不过这无关大局。重要的是在经过一番努力之后，我们终于发现，很难用形式符号的运算来判定某一个数是不是素数。由此可见，形式运算把握某种概念的能力是多么有限。

也许从埃舍尔的《镶嵌图案》（图6）中我们可以获得某种启示。这是一幅图形和背景天衣无缝地配合在一起的图画。如果我们把白色的形象视为图形，那么黑色的形象就构成了它们的背景。反之，如果我们把黑色的形象视为图形，那么白色的形象就构成了背景。它们互相补充，共同组成一个整体。这使我们联想到，素数与合数也是以互补的方式构成统一的整体。

显然，识别合数要比识别素数来得容易。因此，我们很自然地想到先来构成一个合数为定理的形式系统，这一点是比较容易做到的。模仿前面的pq系统，我们来构造一个tq系统。这里的t可以理解为乘法，x、y、z则分别表示连字符号的串。xtyqz为tq系统中的一条定理，当且仅当x乘y等于z。这种系统也和pq系统一样简单，它只有一条公理和一条推理规则：公理：xt－qx为定理，x为连字符号串。

规则：如果x、y、z为连字符号串，xtyqz为定

理，那么xty－qzx也是定理。

有兴趣的读者可以模仿pq系统的推理，根据这条公理和规则构造该系统的定理树的一部分。看看是否能用这种方式把握乘法的概念。

当然，乘法是比加法更为复杂的概念，但是我们现在却用形式符号及其运算将它控制住了，就象埃舍尔在《解放》这幅画中控制飞鸟一样。那么对于素数这个概念又怎样呢？

看来有一个很机智的方案，这就是运用已经构造好的tq系统定义一个形式为Cx的定理集合。当x代表的连字符号串数目为合数时，那么Cx为其中的一个元。我们给出以下的规则：

规则：如果x、y、z是连字符号串，如果

x－ty－qz是一条定理，那么Cz也是一条定理

这是合数定义的形式化。这里之所以用x－y－是避免出现连字符号串为1的情形。这样我们就可以用形式为Cx的定理集合表示合数的集合，用形式为Px的定理集合表示素数的集合。C型定理与P型定理的相互关系是不言而喻的：

规则：如果x为连字符号串，那么Cx不是定理

Px就是定理，反之亦然。

不过，我们不要高兴得太早了，要确定一个连

字符号串不是C型定理并不是一种明显的形式运算。就象我们前面提到过的数学游戏，要想确定MU不是MIU系统中的一条定理就必须跳出这个系统。上面提出的这条规则也是这样，它与我们前面提到的几种形式系统中的规则不同，它要进行的运算是非正常的运算，即在系统之外的运算。这就是说它要浏览一张“非定理”的表。而要生成这样一张表就要在系统之外进行运算。现在的问题是，这个系统的所有定理按照共同的规则运算具有共同的形式，那么所有的非定理是否也具有共同的形式呢?这一点并不是理所当然的。因为系统的定理可以从正面给出定义，而非定理却是从反面来定义的。

我们再来看看斯科特的《图案画》（图28）。我们可以用两种方式来描述黑色区域的特点：

1.作为白色图形的背景。

2.将白色区域平移后涂上黑色。

在这幅图中，用两种方式得到的结果是一样的。这样的性质在形式系统中就表现为一个形式系统中的非定理可以成为另一个形式系统中的正定理。在音乐中，与此相类似的是主题与伴奏。主题是引人注目的，而伴奏在某种意义上讲是隐藏在后面的。在一般的乐曲中，伴奏并没有什么明显的旋

律。但是在巴赫的一些作品中，伴奏也有旋律。它们起着“图形”的作用。从这种意义上讲，巴赫的这些作品是“互衬”的。

现在让我们回到形式系统的正定理与非定理的问题上来。在我们的例子中，C型定理是正空间，P型定理是负空间。我们已经找到一种方法可以用符号把素数表示成某个形式系统的非定理集合，即负空间。现在能否再将其表示成正空间呢？

按照人们的直觉也许会认为这没有什么本质性的困难。为什么图形和背景就不能携带同样多的信息呢？然而无情的事实恰恰是：

存在着这样的形式系统，它的负空间（非定理的集合）不能成为任何形式系统的正空间（定理的集合）。

这个结果在深度上与哥德尔定理相当。把它表述成数学术语就是：

存在着递归可枚举的集合，它是不可递归的。

递归可枚举相当于可以用背景来“反衬”图形。递归相当于图形与背景可以“互衬”。这就是说递归的集合不但本身是递归可枚举的，而且它的补集也是递归可枚举的。

这里有必要说明一下递归可枚举的概念。所谓

递归可枚举的集合就是可以从这个集合的一组基元出发，重复地运用运算的规则来生成集合中的每一个元。这是递归的本质，即用简单的说明来定义一些东西。我们提到过的费波那奇级数和罗卡斯级数就是递归可枚举集合的出色例子。

根据上面的结果（我们暂时先承认它是正确的）就可以得出下面的结果：

存在着这样的形式系统，没有适合于它的形式判定过程。

为什么会有这样的结果呢？所谓判定过程就是识别定理和非定理的方法。如果一个形式系统的定理集合是递归可枚举的，它的补集——非定理的集合也是递归可枚举的，那么我们就可以把所有的定理和非定理都列成表。通过与这张表的对照就可以判定任何一条陈述是定理还是非定理。但是如果一个形式系统的定理集合是递归可枚举的，可是非定理集合却不是递归可枚举的，那么就无法做到这一点。因此并非所有的形式系统都有形式的判定过程。这个问题也是与自然数的有序性问题深刻联系在一起的。被我们认为是非常有序的自然数的数列总会表现出某种无序性来。埃舍尔的画《有序和无序》（图18）巧妙地表现了两者之间的关系。

5

哥德尔定理

5.1 原始递归与递归

前面我们已经用图形和背景形象地比喻形式系统中的定理与非定理，叙述了一个与哥德尔定理具有同样深度的命题。即存在着递归可枚举的集合不是递归的。递归的概念在哥德尔定理的证明过程中起着极为重要的作用。因此我们进一步用计算机科学的语言来描述原始递归与递归的概念。为此我们在这里引进3种计算机语言，即Bloop、Floop、Gloop语言。这三个词很象是一组绕口令，又象是

沉船激起的水声。我们希望读者细心注意它们的定义，因为这是我们进行论述的基础。

我们仍然在形式数论系统中进行讨论。这是一个能够反映自身的系统，或者说具有“自我认识”能力的系统。这种自我认识具有一个很突出的特点，它和自然数的有序及无序的问题联系在一起。我们已经看到，形式系统中可以有一些规则使系统中的串增长或缩短。这就可能使对这些串的探索漫无止境。哥德尔编码的发现表明，对于具有某种性质的串的探索都有算术上的对应物，即与对于具有相应算术性质的整数的探索同构。因此对于形式系统判定程序的探索也包括解开这样的秘密，为什么会有无限长的探索。这就是说，为什么整数总有某种无序性。

有些整数序列一看就知道有什么规律性，有时却要通过某种抽象的滤波机制才能发现，即似乎无序的序列后面的结构。但是怎么能保证总能找到这样的结构呢？

我们的目的是要讨论对具有各种性质的自然数的探索。为了描述这种探索的长度，我们先要定义一些基本步骤。我们可以把这样一些步骤看成是基本的：

把两个自然数相加

把两个自然数相乘

确定两个自然数是否相等

确定两个自然数中哪一个更大（或更小）

如果我们想用这些基本步骤来描述一种试验方法，那么就要包括一种控制结构。这就是描述进行各个步骤的顺序，什么时候返回来重复进行某些尝试，什么时候可以跳过一些步骤，什么时候应该停止。任何一种描述如何完成某项任务的算法都包括这样两部分：(1) 所要完成的操作；(2)控制指令。因此我们要想搞出一种有预先规定长度的计算语言，也必须具有原始的控制结构。实际上Bloop语言的特点就是有有限组的控制结构。它不容许无限地扩充步骤，也不容许无限地重复一组步骤。Bloop语言中唯一的控制结构就是有界的圈。这就是说可以重复执行的指令，重复的次数不能超过某个最大值，这个值就是它的上界。

这种程序中所有上界的值并不是明确给定的。但是可以在进入这个圈之前计算出来。例如在计算2^{3^n}时，有两个圈。先计算 3^n 的值，其中包括n次乘法。然后再计算 2 的幂，其中包括 3^n 次乘法。因此第二个圈的上界是第一次计算得到的结果。

在计算机程序中，有一部分程序可以组成一个单元，它们作为一个整体而被调用。我们把这称为程序的块化，所形成的单元就是程序块。这些块可以连接起来而形成总的程序。

Bloop语言有两个特点。第一个特点是，一旦定义了某个步骤后就可以从内部引出后面步骤的定义。这样一来，在程序中所定义的操作就和那些基本步骤一样简单。也就是说Bloop语言可以自动地形成块。第二个特点就是操作过程的输出可以不是数值而是是或者非。这样的过程更确切地讲是一种检验而不是函数。

Bloop语言具有无引入地址的结构。这种结构使得程序具有自备的性质，其中每一步可以引出下一步的定义。我们把能够用Bloop语言的程序计算的函数称为原始递归函数。而能用Bloop试验进行鉴定的性质称为原始递归性质。例如2^{3^n}是一种原始递归函数。“n为素数”则是一种原始递归性质。

为了说明Bloop语言与形式数论系统之间的关系。我们先来区别一下“可表达性”与“可描述性”。在形式数论系统中“表达”一个命题，仅仅是把普通语言陈述的命题译成相应的形式符号。这和该命题是否成为系统中的定理无关。而“描述”

一个命题则是更强的概念。它若为真便是定理，若为假则是非定理。

形式数论系统可以表达所有的数论中的命题。而“在TNT中可以描述什么样的性质？”这个问题就是说“TNT是具有什么能力的公理系统？”如果能够在TNT系统中描述所有的数论命题，那么TNT系统就能回答数论中的所有问题。它对于数论中的命题就是完备的。

至少可以肯定一点，对于原始递归的谓词TNT是完备的。这就是说，如果对于自然数的某种性质可以写出Bloop试验程序，那么这种性质就可以在TNT内描述。

我们现在设想这样一个很怪的概念：所有Bloop程序的圈。这样的圈当然是无限的。我们再考虑它的一个子圈。它是通过连续的3种滤波操作而得到的。经第一步滤波后只剩下无引入地址的程序。第二步滤波则把所有的检验排除掉，只剩下那些函数。第三步滤波后只剩下有一个输入参数的那种函数。最后剩下的子圈是什么呢？是计算只有一个输入参数的函数的所有无引入地址的Bloop程序构成的完备圈。我们把这种特殊的Bloop程序称为蓝程序。

我们对每种蓝程序给定一个指数，可以把它们按照长度加以排列。当然可以有许多不同的蓝程序具有相同的长度。因而我们采用字母表的顺序加以排列。这里所说的字母表顺序是广义的，可以用字母分别代表Bloop程序中的所有特点。

运用康托尔的对角线法，我们总能找到这样一类蓝程序，可以用它来定义一个新函数，它有明确的定义，是单变量的函数，但却无法用 Bloop语言来编制程序。这就是说，用这种蓝程序定义的函数不在原始递归函数之列。我们把这种函数称为蓝对角函数。

用康托尔的对角线法可以证明，如果将所有的实数进行编目，那么总有一些实数不在其中。这就是说对实数进行完备的编目是做不到的。我们只考虑从 0 到 1 之间的实数。假设所有这样的实数可以列成一张表。因为从 0 到 1 的实数总可以表示成无限的小数，因此我们可以设想把 0 到 1 之间的所有实数编上号而得到这样一张表：

r（1） 0．1 4 1 5 9 2 6 5 3……

r（2） 0．3 3 3 3 3 3 3 3 3……

r（3） 0．7 1 8 2 8 1 8 2 8……

r（4） 0．4 1 4 2 1 3 5 6 2……

我们依次取这张表上对角线上的数字作为新的实数相应各位上的数字，得到一个新的数字 0．1 3 8 2……然后构造数d，使d任何一位的数字都与 0．1 3 8 2……相应位上的数字不同。显然这个数d是 0 到 1 之间的实数，但并不在这张表上，因为它与表上任何一个实数都不相同。由此可以证明从 0 到 1 的实数不可能与自然数完全一一对应。

康托尔对角线法的本质在于以两种方式使用整数，也可以说是在两种不同的层次上使用整数。一方面作为行的指标，另一方面又作为列的指标。

也许有人会想，如果把d补充到表上去，这张表不就更完备了吗？其实这是无济于事的，因为对角线法同样适用于添加d后的新表，从而可以得到不在新表上的d′。这个过程可以重复进行下去，这就好象乌龟的唱片不断使蟹的唱机粉碎一样。

我们已经用Bloop语言写成的程序定义了一类原始递归函数和自然数的原始递归性质。我们也看到，Bloop语言无法描述我们可以用语言定义的所有自然数的函数。我们用康托尔对角线法构造的蓝

对角函数就无法用Bloop语言来描述。

定义Bloop语言的性质可以理解为它的圈是有界的。如果我们去掉这种限制就可以得到一种新的计算机语言Floop（意思是含有自由圈）。

仿照前面的蓝程序的定义，我们可以构造一种绿程序：

所有计算只有一个输入参数的函数的无引入地址Floop程序的完备圈。

在Floop程序中可以包括无界的圈，它可以是无终止的程序。如果排除这种无终止性，我们就可以得到所谓的红程序：

所有计算只有一个输入参数的函数的无引入地址的Floop程序，而且这些程序对于它们的所有输入都是可终止的，这些程序构成的完备圈。

我们把无法用Floop语言描述的程序称为Gloop程序。

现在我们已经有好几种计算机语言。可以严格地证明，那些语言具有和Floop语言同样的能力。因此，一般都认为不可能有比Floop语言更强能力的计算机语言。著名的丘奇—图林命题就表达了这种思想：

人能够计算的也是机器可以计算的；

机器可计算的也是Floop程序可计算的。

如果我们把Bloop程序可计算的函数称为原始递归的，那么Floop程序可计算的函数就可以分为两类：（1）可以用有限长的Floop程序计算的函数为一般递归函数。（2）只能用无限长的Floop程序计算的函数为部分递归函数。人们通常所说的递归函数是指一般递归函数。

有趣的是TNT不仅可以描述所有的原始递归函数，而且可以描述所有的一般递归函数。我们不能在这里证明这两个结论。这不是本书的目的。我们的目的是要说明形式数论系统的不完备性。

5.2 哥德尔编码与自我反省

上一节我们引进了Bloop、Floop、Gloop语言的定义，用这些语言描述了原始递归与递归的概念。现在我就可以介绍TNT中形式上不可判定的命题及有关的形式系统了。

哥德尔在1931年发表的论文专业性很强。我们不可能在这儿介绍它的数学细节。我们要介绍的是完成主要定理证明过程的两个思想。第一个思想就是找到了一种方法可以把TNT中的串解释成说明

TNT中的其他串。这就是靠哥德尔编码。第二个关键思想就是这种自我反省的情况可以集中在同一个串上，也就是TNT中的一个串可以解释成在说明自身。而这种方法从本质上讲可以归结为康托尔的对角线法。

我认为，如果一个人对哥德尔定理真正感兴趣，就会认识到它的证明实质上是融汇了上述的两个思想。这两个思想是极为出色的，而把它们结合在一起就无疑是一种天才了。我个人则认为第一个思想更为重要。正是哥德尔编码涉及到符号运算系统的意义与相互关系。

明确了这两个基本思想，我们进一步说明定理的证明过程。有了哥德尔编码，我们可以在TNT系统中定义“证明对”的概念。证明对是一对以特殊方式联系起来的自然数对。这种思想是：

两个自然数m与n形成TNT中的一个证明对，当且仅当m为TNT中可由哥德尔数n推导出来的哥德尔数。

现在我们指出其基本事实1：

作为证明对的特点就是具有原始递归的特点，因而可以用Bloop程序来检验。和其他的数论性质不同，这种性质的显著特点是定理数。所谓n是一

个定理数就是说存在一个数m可以与n构成一个证明对。

从上面的基本事实1我们可以得出基本事实2：

构成证明对的性质可以用Bloop程序来检验，因而可以在TNT中用两个变量的公式来描述。

前面已经说了，考虑一种性质的“表达”和“描述”之间的区别是很重要的。例如作为TNT的定理数的性质可以表达成：

$$\exists a:\ \text{TNT—证明对}\ \{a,\ a'\}$$

这可以解释为：“a′是TNT中的一个定理数。”但是我们并不能说这个式子“描述”了这个概念，因为我们并不能保证这种性质是原始递归的。事实上，我更倾向于怀疑这一点。这种怀疑被证明是有道理的。作为TNT中定理数的性质并不是原始递归的，但作为证明对的性质却是原始递归的。后者既是可以表达的，也是可以描述的。

通过哥德尔编码，我们可以把上面这个式子与TNT中的一个数对应起来。这样TNT中的数就可以描述TNT中数的性质。这就实现了TNT中定理的反省。这是证明第一部分的实质。

我们现在进一步介绍证明中第二个重要思想。

那就是把这种反省集中到同一个式子里。为此我们需要考察一下当我们以一种简单的方式改进式子的结构时，哥德尔数将发生什么变化。实际上，我们考虑这种特殊的改进：

用特定的数代替所有的自由变量。

请看下面的例子：

式子	哥德尔编码 → 哥德尔数
a = a	262 111 262
用数字 2 代替所有自由变量	↓
SS0 = SS0	123 123 666 111 123 123 666
. . .	. . .
~∃a:∃a′:a″ = (SSa· SSa′)	223 333 262 636 333 262 163 636
用数字 4 代替所有自由变量	262 163 163 111 362 123 123 262 236 123 123 262 163 323
	↓
~∃a:∃a′:SSSS0 = (SSa· SSa′)	223 333 262 636 333 262 163 636 123 123 123 123 666 111 363 123 123 262 236 123 123 262 163 323

上面的左右两行是同构的。在右边一行中，从一个哥德尔数到另一个哥德尔数之间的函数关系不难用算术运算的术语来描述，不过在这里没有必要这样做。重要的是我们知道在原来的数，插进去的数以及最后得出的数之间的关系是原始递归的。也就是说，可以写出一种Bloop程序，把三个数字输入后可以得出是或非的回答。如果答案是肯定的，我们就可以说，在它们之间有“替代关系”可以用下面的式子来描述：

SUB{a, a′ a″}（SUB是“替代”的意思）

在我们给出的第一个例子中，下面的式子就是TNT中的一条定理。

SUB {SSS……SSSO/a, SSO/a′, SSS……SSO/a″}

262 111 266个s 123 123 666 111 123 123 666个s

而下面的式子显然不是一条定理：

SUB{SSSO/a, SSO/a′, SO/a″}

有了这些基础，我们就可以把它们综合起来构造一个句子，它的意义可以解释成：“TNT中的这个串不是TNT中的一条定理。”显然这是一个怪圈，是和爱皮梅尼特悖论同构的。做到这一步就万事大吉了。不过我们不能高兴得太早了。虽然一切

条件都已经具备，但是要在实际上做到这一点仍然不是一件轻而易举的事。要在TNT中表述这个句子还需要一个举足轻重的概念，这就是要运用“奎因”（quine）的技巧。所谓“奎因”一个句子，就是用整个句子代替该句子中的某个成分而构成一个新句子。而这两个句子可以说是同构的。例如：“据我所知并不是一支歌”经奎因而成为“‘据我所知并不是一支歌’据我所知并不是一支歌。”同样，在TNT中我们可以用式子所对应的哥德尔数去代替式子中的自由变量而产生一个新的哥德尔数。我们把这称为“算术奎因”。运用这种方法，我们就可以在TNT中实现自我相关。让我们看这样一个例子。

给定一个式子a = S0，它的哥德尔数为262 111 123 666。让我们把这个数插进式子中去就得到：

$$\underbrace{SS\cdots\cdots S}_{262\ 111\ 123\ 666\text{个}S}0 = S0$$

这个结果显然是错误的。因为让 262 111 123 666 和 1 相等是荒唐的。但是如果我们从一开始就写成～a = S0那么就会得到一种正确的陈述。

如果我们在“替代关系”中把两个自由变量换成相同的。那么就会有SUB{a′′, a′′, a′}。这是以

两种方式使用同一个数字。我们可以写成：

算术奎因{a′′，a′}

这个式子的含义是："a′是这样一个式子的哥德尔数，它是由哥德尔数为a′′的式子通过算术奎因而得到的。"在我们所举的例子中，庞大的数字：

123 123……123 123 666 111 123 666

262 111 123 666个"123"

就是"算术奎因"a＝S0所得到的式子的哥德尔数。

就象我们可以重复奎因的技巧一样，我们也可以重复地使用算术奎因的技巧，就是对于一个描述算术奎因的句子进行算术奎因。于是我们就可以写下这样一个式子：

～∃a：∃a′：＜TNT证明对｛a，a′｝∧算术奎因｛a″，a′｝＞

我们称它是"G的叔叔"，记作"u"。u当然也有它所对应的哥德尔数。我们要做的最后一步就是对G的叔叔"u进行算术奎因。这里的自由变量只有a″，用u的哥德尔数代替它就得到：

～∃a：∃a′：＜TNT证明对｛a，a′｝∧算术奎因｛sss…sso/a″，a′｝＞

u个S

不管你信还是不信，这就是哥德尔串，我们称它为“G”。现在我们马上要回答两个问题：

（1）G的哥德尔数是什么？

（1）G如何解释？

对于第一个问题，我们想想，G是如何得到的？是对u进行算术奎因而得到的。因此G的哥德尔数就是u的哥德尔数进行算术奎因而得到的数。这是一种简化的说法，整个过程前面已经讲了。

对于第二个问题，我们可以用普通语言来叙述这个式子：

不存在数a和a′，便得下面两点同时成立（1）它们形成TNT中的一个证明对，（2）a′是u经算术奎因所得到的哥德尔数。但是算术奎因u所得到的哥德尔数确实是存在的。因此上面这句话可以表述成：

不存在数可以与算术奎因u的哥德尔数形成证明对。

现在你可以看到会有什么结果了。

一个式子的哥德尔数是算术奎因u得到的哥德尔数，那么这个式子不是TNT中的定理。然而这个式子恰恰就是G本身，因此这就是说：

G不是TNT中的定理。

这样我们就从较低层次的解释——数论中的句子逐渐推导出较高层次上的解释——元TNT中的句子。这里的主要结果可以重申如下：

如果G是TNT的定理，那么它一定为真。但是G说明了什么呢？说明了它自己不是一条定理。于是从它是定理得出它不是定理：构成了矛盾。

如果G不是TNT的定理又怎样呢？这是可以接受的，它不会导致矛盾。但G不是定理是由G说明的，因此G是真的。又因为G不是定理，这就是说有一条真命题不是TNT中的定理。于是我们就在TNT中发现了一个漏洞——存在着一个不可判定的命题。

哥德尔找到了一种明确的方式用TNT中的式子来陈述"TNT是一致的"然后又证明这样的陈述要成为TNT中的一条定理只有在"TNT是不一致"的条件下才有可能。这是个看来不合情理的结果，却对它进行了严格的数学证明。其结果沉重地打击了那些致力于寻找数学无矛盾性严格证明的乐观主义者。

哥德尔定理证明了TNT的不完备性。这是前面讲过的那种ω型不完备性。就是说，有一族塔型的串，它们有无数多个。每一个串都是定理，但是

“概括”它们全体的全称命题却不是定理。这很容易用例子来加以说明：

$\forall a$：～$\exists a'$：＜TNT证明对｛a，a'｝入算术奎纳{SSS…SSSO／a''，a'}

u个S

这个串不是一条定理，但我们可以证明，与此有关的塔形族中每一条都是定理。因为它们每一条的意思分别是：

“0和算术奎因u的哥德尔数不构成 TNT 的证明对”

“1 ”

“2 ”

等等。

因为G不是定理，因此没有一个整数与G的哥德尔数形成证明对。由此可见这一族中每一条都是定理。这就证明了TNT是ω不完备的。

5.3 无法弥补的漏洞

我在在TNT中发现了漏洞。这个漏洞有没有办法补上呢？人们可能会联想到几何学的情形。由4条几何公理组成的绝对几何学可以向两个方向扩充，即欧氏几何与非欧几何。这就启示我们可以把G或

～G作为一条新的公理而加到原来的系统上去。我们就以把G作为新公理为例，把它加到TNT中构成新的系统为TNT＋G。

TNT的漏洞可以表示成："我无法在TNT中得到证明。"

而在TNT＋G中存在着同样的漏洞，可以表述成"我无法在TNT＋G中得到证明。"

因此，不管怎样填补漏洞、扩充形式系统，新的形式系统仍然具有哥德尔的结构。

这个原理可以用康托尔的对角线法来很好地说明。康托尔用这种方法来"发现""实数完备表"中遗漏的实数。我们可以用同样的方法来"发现"形式系统中的漏洞。

(1a) 给定表L，构造它的对角线数d。

(1b) 把d加到表L里构成新表L＋d。

(2a) 给定表L＋d，构造它的对角线数d′。

(2a) 把d′加到表L＋d里构成新表L＋d＋d′。

⋮

显然用这种方法是无法构造所有实数的完备表的。同样用补充公理的办法弥补漏洞也是徒劳无益的。

需要说明的是，一个形式系统要出现这种哥德

尔式的不完备漏洞必须满足 3 个条件：

1.这个系统足够完备使得所有有关数的陈述都能在系统内表达出来。

2.所有的一般递归关系都能在系统内用公式表达出来。

3.公理和那些规则所定义的符号模式都能通过某种有限的判定程序来识别。

凡是满足这 3 个条件的一致的形式系统，哥德尔理论都是适用的。

按照卢卡斯的观点，哥德尔理论恰恰证明了机械论的破产，这就是说，不能把思维解释成机械。

他的推理过程如下。如果有一架和人一样聪明的计算机就可以完成人所能完成的任何智力活动。让我们考虑任何一种特定的形式系统。不难写出计算机的程序，它可以系统地产生该系统的各条定理。这种程序由两部分组成。第一部分子程序可以打印出公理，给出公理系统的模式。第二部分子程序则根据已知的定理（当然也包括公理）运用推理规则产生新的定理。

我们可以拟人地说，这一程序“懂得”关于数论的某些事实。也就是说它懂得打印出来的那些事

实。如果它无法打印出数论的某些真实结果，那么它当然是不懂这些结果的。如果能够证明，人懂得一些计算机程序所不懂的东西，那么就说明人要更胜一筹。

于是卢卡斯指出，人可以把哥德尔的技巧用于任何一种形式系统，我们总能比它本身懂得更多一点。虽然看起来这是有关形式系统的论断，但是只要稍加改进就可以用来证明，不可能产生和人同等水平的人工智能。

我们还应该更深入一步了解，为什么卢卡斯说，计算机的程序不管怎样设计都不可能比我们知道得更多。其基本思想就在于，我们总是在系统之外，这样就可以完成“哥德尔化”的操作。

为了使我们对此有一种直观的印象，我们来看看埃舍尔所画的《龙》（图19）它最显著的特点就是所表现的主题——一条龙咬住自己的尾巴，其中包含了哥德尔定理的意义。但是还有更深一层的意思。埃舍尔自己对此作了一番说明：

1.我们的3维空间是我们所知道的唯一真实的存在。2维空间则和4维空间一样虚幻，因为没有一样东西是绝对平的，那怕是最精细地磨平的镜子。但是我们仍然容许这样的约定，墙和纸是平

的，甚至还要在这样的平面上来表现空间的假象……

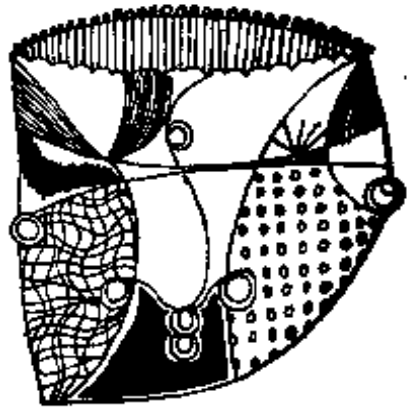
2.但是这条龙是多么想成为空间性的，虽然它仍然是完全在平面上。在画着龙的纸上有两个剖面。它以这种方式折迭，在画面上留出两个正方形的开口。这条龙是顽强的，尽管它是2维的，仍然设想自己是3维的，于是它的头从一个洞里伸出而尾巴则通过另一个洞伸出来。

我们可以随心所欲地折腾埃舍尔的画，把它从书上撕下来，折迭起来、挖个洞再穿过去最后把这一堆乱七八糟的东西照下来，于是就又成为2维的了。

如果我们深思一番人是否能超脱自己——或者计算机程序能否超脱自己，这是非常有趣的。确实，一种程序可以改进自己，但是这种改进能力却是在程序一开始时就已经规定好了，因此并不能把它看成是“跳出该系统”的例子。TNT系统可以反省自己，但是无法跳出自己。一种计算机程序可以改进自己，但是无法改变自己的指令，而只能按照那些指令来改变自身的某些部分。因此跳出系统的愿望往往只是一个无法成为现实的梦。

6

大脑与思维



6.1 描述的层次

哥德尔理论是人类思想最深刻的成果之一。它也必然会影响到人们对于自身的认识，特别是对于思维过程及其物质基础——人的大脑机制的认识。

通过思维过程人认识了世界，并且用语言描述这个世界。我们可以说思维过程具有把握现实世界意义的能力。但是我们已经看到，无论是形式系统的符号、埃舍尔的画还是巴赫的赋格都可以从不同的层次去理解，从而获得不同的意义。

这种现象有时会使我们迷惑不解，有时却又显得非常自然。例如我们从电视屏幕上看到一连串的画面，我们明明知道这不过是显象管的电子枪在萤光屏上激发的一组亮点。但是我们仍然会被这些画面所表现的情景所吸引住。一般人所看到的只是画面的清晰度、色彩、亮度等等。只有电视机的专家或业余爱好者才会想到电子线路和电子元件的性能。这就是说，我们可以有截然不同层次的理解却又不会造成混乱。音乐也是这样，同样的乐曲可以使一部分听众感到恐惧，却会使另外一部分听众感到振奋，这取决于听众的个人情感、心理和气质。再拿埃舍尔的画《带魔带的立方架》（图17）来说，魔带上的小泡究竟是凹的还是凸的呢？一会儿它们的一部分凹进去，另一部分却凸了出来，一会儿凹进去的那部分凸了出来，突出来的那部分却凹进去了，就象在变魔术一样。在心理学的实验中还有许多这样奇妙的图。读者也许不会忘记，前面提到过的《凹与凸》也有同样的效果。这就是说画面的视觉效果取决于你用什么方式去理解它。巴赫的赋格曲也是如此，你可以单独欣赏其中的一个音部，也可以欣赏它们的总体效果。它往往会使听众感到迷惑就是因为听众不知不觉地一会儿采取前一

种方式，一会儿又采取后一种方式。

人工智能的研究是与研究人类的思维相辅相成的。对于人类思维过程的研究推动了人工智能研究的发展，促使电子计算机模仿越来越多的智力行为。反过来，人工智能的研究也对我们理解思维的过程以有益的启示。

人工智能所要解决的核心问题之一就是如何才能填平两种不同层次之间的鸿沟，如何才能构造一个系统既能接受某一个层次的描述又能产生另一个层次的描述。我们仍然以计算机下棋为例来进行分析。

早在20世纪50年代，人们就已经相信计算机可以比象棋大师更好地估价每一步棋。计算机估价各步棋的容量和速度远远超过一位最出色的棋手。由此推论，计算机要战胜象棋大师是轻而易举的事。然而实际上并不是这样。棋手和计算机较量的结果表明，一位出色的棋手可以有把握地击败最先进的下棋计算机程序。其主要理由就是因为棋手不仅可以根据每一步棋去分析局势，而且可以从更高的层次去理解。而这一点，他的对手却做不到。对象棋的研究还表明，象棋大师对棋局的理解往往形成块。也就是说，象棋大师头脑中的棋局不是以每个

子的位置组合而成的，而是以一部分子的相互关系形成的块组合而成的。因此象棋大师只需10秒钟就能把奕局记下来，而新手在同样的时间内却只能记住一半。但是如果把棋子杂乱无章地放在棋盘上，那么要记下这样的“局势”，象棋大师并不比新手来得高明。在考虑下棋的策略时也是这样，象棋大师并不见得比新手考虑更多的步骤。一般来讲，他只考虑很少的几种着数。他不会在分析棋势时去想到那些容易失败的着数。其实初学者只要经过一段时间的实践就会发现，有一些着数是万万不可取的。一旦他明白了这些，他的下棋水平就提高了层次。而象棋大师则把这种技术发展到更高的水平上来。

同样，天才的数学家一般不会去考虑那些普通数学家也许会考虑的途径。他们似乎有一种本能会“嗅”出最有希望的途径，并且毫不犹豫地选中了它们。也许有一天，计算机依靠多看几步的能力可以战胜棋手，也能在数学证明中表现出令人望尘莫及的能力。但是仍然不会改变这样的结论，智能的关键问题是要创造更高层次的描述。

既然人们可以从不同的层次去理解，那就难免要造成某种混乱。人的精神结构是多层次的，也是一个尚未被人们充分理解的系统。对于人们各种行

为的解释有许多相互竞争的理论。例如当人们谈到心理上的“动力”时，有的把它归因于性欲，有的归因于权力，有的归因于荣誉，有的归因于爱情等等。但是人们并不清楚这些动力来自于人类精神结构的哪个层次。对于这样一个多层次的系统，我们自己又从不同的层次去理解，那么所造成的混乱也就可想而知了。

计算机也是一个具有多层次的系统。就拿计算机程序来讲，有机器语言、汇编语言、编译程序等等。但是这些不同层次的程序又都是用计算机科学的语言写成的，因而各个层次的不同描述在某种意义上讲又是彼此相似的。

机器语言可以说是计算机的基本语言。计算机的硬件只能处理这种语言。典型的机器语言包括这样的指令：

地址、打印和转移。

按照这些指令，计算机的硬件不断运行，使计算机从一种状态转移到另一种状态。

比机器语言更高一层的是汇编语言。在机器语言与汇编语言之间并没有什么明显的界线。我们可以说汇编语言是由机器语言形成的块，这些块可以作为单元使用从而提高计算机的效率。

更高一级的层次是编译程序语言。这种语言使计算机能将高一层次的语言译成低一层次的语言。当然编译程序语言也是用计算机语言写下来的。一旦写下了某个编译程序的核心部分，它就可以写出更大的编译程序来，就能把更大的编译程序译成机器语言，使计算机硬件能够进行处理。这就是所谓的自展过程。自展过程好象一个具有一定语言能力的孩子一样，他能运用已经掌握的词汇和语法去获得新的语言知识，这就使他的词汇量和语言的流畅程度可以突飞猛进。

人们还发明了翻译器，它能将高层次的计算机语言译成机器语言。所不同的是，这种翻译器可以译一行就执行一行。这样使用翻译器的人就不必写下全部程序，而可以一行一行地设计。想一行、让机器执行一行，检验一行，然后再往下想。显然，这种方法给设计者带来极大的便利，便于设计也便于修改。因此它深受人工智能研究者的欢迎。

计算机的多层次结构对于使用者来讲是十分必要的。他可以不必去考虑那些与他无关的低层次。这就好象一位乘坐飞机的旅客并不需要去关心飞机油箱中的燃料，也不需要去考虑当时的风速或者机上能供应多少份午餐鸡。这些问题应该留给航空公司

的各级工作人员去考虑。只有当系统的运转出现某种故障时——例如旅客的行李没有按时到达——他才会意识到在这个系统的某些低层次上出了毛病。

计算机的重大进展就是向更高的层次发展。因为发展计算机的目标之一就是要更接近人的自然语言。人们的自然语言中包含着许多含糊而有弹性的地方。在自然语言中往往使用不完整的句子、变形的意义和不合语法的描述。然而这种“不规范”的语言却能传递人们想要传递的信息，表达人们想要表达的感情。计算机所用的程序语言却必须百分之百地遵守规则。那么如何才能使人和机器对话呢？这就要发展更高层次的计算机语言，使它既便于人们使用又满足计算机所要求的不含糊性和精确性。

现代人工智能研究的主要领域之一就是自动编制程序。这就是发展一种高层次的语言。这种高级精巧的语言具有这样一些功能：推广例子；纠正计算机的打印和语法错误；能够理解那些含糊其辞的描述；能够运用模型提供预言性的猜测；当什么地方不清楚时还能提出问题。实现这些希望的努力是在可靠性和灵活性之间走钢丝。这种技能是高超的，却并非是不可能的。

计算机最明显的层次表现在硬件和软件的差别

上。这是程序和计算机部件之间的差别，是一系列指令与执行这些指令的元件之间的差别。如果说一架钢琴或者一架电唱机是硬件，那么乐谱就是软件。一架电话是硬件，电话号码就是软件。我们人类也有硬件和软件。我们必须承认人在生理上的那种固定性，我们无法补救自己的一些生理缺陷，也无法按照自己的爱好长出某种颜色的头发来。但是我们有能力改变自己的思维模式，用新的概念网络进行思维操作。我们无法使神经元的兴奋加速或者减慢，无法改变大脑各部分与神经系统其他部分的联结方式，无法重新设计整个神经系统的内部结构。也就是说我们无法对自己身上的这些硬件进行选择，但是我们却能控制如何进行思维的过程。

当然思维的有些方面也是我们无法控制的。我们无法使自己变得更加机敏，无法按照自己的愿望在很短的时间内掌握一门外语，如此等等。

一般来讲，多层次的系统可以分成两类。有一类系统的某一层次的行为可以掩蔽另一层次上的行为。不管在低一层次上发生什么事件都不会在高一层次表现出相应的变化。典型的例子就是容器里的气体系统。尽管在微观的层次上分子运动激烈，可是在宏观的层次上，这是一个具有一定温度、压力某

和体积的系统，是平静而稳定的系统。但是也有象弹球游戏装置这样的系统，低层次上的单独事件也会在高层次上反映出来。球撞击柱子的角度、撞击力的大小都会影响球下滚的途径，并且影响到最后的结果。

计算机则把这两种系统的不同特点结合起来了。它既包括这样的元件，它的行为遵循统计的规律性，是可以在宏观上预测的。但是也包括这样的元件，在程序中的每一个比特的变化都会影响到最后打印出来的结果。

由“可靠”的子系统所构成的系统在我们日常生活中也起着不可估量的作用，它们可以说是稳定性的支柱。这种系统的模型可以是完全确定的。或者一个系统的宏观量是以一种我们观察不到的方式取决于内部的微观参数。还有一种有趣的现象。比如有一位运动健将以10秒跑完百米的全程。可是10秒这个数据与身体的哪一部分直接相关呢？问题的关键不在这里。因为10秒是他的身体结构、心理状态、行为方式以及当他跑步时无数其他因素共同作用的结果。这种结果是可以重复的，但是无法描述产生这种结果的规律。人们的轻信也是这样一种现象。如果有人认为可以用一种神经系统的手术来弥

补自己过于轻信的缺陷，那他确实是太轻信了。

6.2 大脑

为了要用机器来代替人的思维功能，人们发明了计算机。先是手摇计算机，然后是电子计算机。电子计算机的模型使人们更加深入地理解自己大脑的机制。这种机制是人类思维过程的物质基础。不过在这种研究基础上提出的大脑机制究竟在多大程度上是合理的，这是我们应该慎重对待的问题。

人脑可以分为大脑、小脑和下丘脑等。大脑的两半球为大脑皮层所覆盖，在那里执行着思维的主要功能。大脑中的主要细胞是神经细胞。每个神经细胞都有突触和轴突，突触接受输入，而轴突则执行输出的功能。无论输入和输出都采用生物电流的方式。神经细胞可以处于兴奋或者抑制的状态。这种行为方式是非常简单的。但是这些简单的元件可以构成极为复杂的网络，从而在更高的层次上把握概念、执行其他思维功能。

神经学家的研究表明，动物的记忆能力并不限于某些部位。但是另外一些实验却表明，人的一些

记忆集中在一些非常小的区域。这是两种看来截然相反的结果，我们可以这样解释，记忆最初是在神经系统的某些特定部位进行编码。但是后来又在大脑层的其他部位重新编码。这很可能是进化的产物。还有一种解释是，这种系统很象是现代化的电话网络。当某一部分发生故障或者被破坏时，并不会使电路中断。因为通路可以绕过这些出毛病的区域。这就是说电话之间的通路并没有固定的途径。但是相互联系的电话又是完全确定的点。从这种意义上讲，这种联系的位置又是确定的。

对视觉过程的研究表明，一个人要辨认出某种形象要有一个“结晶”的过程。这并不单纯取决于对视网膜上光敏细胞的物理刺激，也取决于大脑的思维模式。一个形象的“结晶”是按一定方式对视觉信号进行处理后才完成的。

撇开具体的神经操作过程不谈，我们可以用符号的网络模型来解释神经系统的行为，描述大脑的状态。不过和一般的形式系统不同的是，大脑所使用的是具有主动性的符号，这些符号所遵循的规则往往和符号混合在一起。在形式系统中，符号与实体的对应关系是固定的，而我们头脑中关于同一实体的描述却是不确定的。大脑思维可以以一种灵活

的、内涵的方式来反映现实的世界。

使用符号是思维过程中高层次的描述，符号可以通过信号的传递而相互联系。这种引发模式很象是在我们所生活的世界里发生的宏观事件。或者说能够在一个与我们的世界相类似的世界里发生。从本质上讲，使用符号的过程与现实世界是同构的。虽然这种同构的方式是复杂的、巧妙的。

但是，马上就会提出这样的问题。在大脑中可以用符号来表示的概念有没有什么限度。因为在概念中有类和个体的区别，而符号往往具有双重的功能，什么时候起哪一种功能就要看它起作用的前后关系。例如化学符号O既可以代表氧元素，是所有氧原子的类，也可以代表一个氧元子。在元素周期表中它执行第一种功能，在表示一个水分子合成的反应式中，它又起第二重作用。

虽然我们也可以按照普遍性和特殊性来把符号分成不同的层次。但是关于“类”的思想是非常广泛而抽象的，是难以用这种方式来加以限制的。其根本原因就是因为我们的思维实际上遵循这样一条原则，最特殊的例子也可以作为一类事件最一般性的代表。既然如此，那么头脑中的符号究竟代表类呢还是代表类的具体例子？是有些符号代表类，另

一些符号代表例子呢，还是每个符号都能起双重作用？

要回答这样的问题是非常困难的。因为一方面，符号代表的具体对象具有它们所属那个类的许多性质。但是它又可以逐渐脱离原来所属的类而成为另外一个类的具体代表。

那么还有一个问题。一个符号能否分离出来呢？就是说割断它与其他符号的相互联系而单独提取出来呢？大概是不行的。因为一个符号总是与其他符号联系在一起而存在、而起作用的。一个符号区别于其他符号的独特性质恰好寓于它与其他符号的相互关系之中。

能够从一个类中抽出具体的实例，或者从具体的实例中抽象出这个类的概念，这是思维的重要基础之一，也是人与动物的不同之处。

既然符号的运用在大脑的行为中起着重要的作用。那么在各个人的大脑之间是否存在着符号水平上的同构呢？这种对应关系不仅表现在符号与符号之间，而且表现在符号的引发模式与引发模式之间。我们在这里所说的符号之间联系是指功能性的关系并不要求一定有物质上的联系。也就是说，人们对符号的运用有没有什么共同的规律性呢？

一方面我们可以看到，要发现人们大脑的软件之间的严格同构关系几乎是没有希望的。但是另一方面我们又可以看到，在一些人的思维方式之间具有惊人的类同之处。这又似乎在表明，他们的一部分软件之间存在着某种同构的关系。这种共同之处特别明显地表现为符号的贮存方式和符号的引发模式上。

我们来比较一下不同语种的符号网络。这种网络很象蜘蛛网。如果我们把这个网络的总体结构称为整体性质，那么在每个结点处相汇的线的平均数可以说是一种局部性质。当我们想判定两个语种的网络是否同构时，究竟应该以哪一种性质为主呢？

不同的语种在词汇、语法和句法方面都有差异。因此这种符号网络也有很大的不同。但是我们仍然希望这种网络的核心部分具有相同的构造，希望有些主要的途径在每一个语种中都是适用的。

但是实际上，无论语言还是思维方式都是深受一个民族的文化的影响。这个民族的历史、地理、宗教、文学、技术水平等等都会影响到这个语种的符号网络。谁要想流利地使用现代希伯莱语就必须熟悉希伯莱语的圣经。

由于语言结构的差异，在翻译时常常会遇到一

些难以找出对应关系词语的问题。这在诗的翻译中表现得最为明显。要把精练的中国古诗译成外文是极为困难的，有时译文简直是令人啼笑皆非的。同样，外国幽默小品中的一些双关语或者文字游戏也很难在中文译文中保留下来。

计算机语言之间的翻译也有类似的困难。我们设想有两个人为两台不同的计算机编制了两套不同的程序。但是两台不同的计算机执行的却是相同的任务。我们现在要比较这两种语言，那么如何进行比较呢？在什么层次上来进行比较呢？假设有一种是编译程序语言，另一种却是机器语言。这样的两种计算机语言能否进行比较呢？当然能。可是怎样进行比较呢？有一种办法就是用翻译器把编译程序语言译成机器语言来进行比较。

但是这样一来也会产生新的问题。因为使用的计算机不同，这两种机器语言可能完全不同。因为这两种计算机采用不同的硬件，因此它们分别适用的机器语言根本无法进行比较。

于是我们只好避开机器语言这个层次而到较高的层次上来。这样一来并不是万事大吉。因为那些成块的语言，即所谓的算法语言，种类繁多、结构也不相同。我们怎样用这些语言写成的程序呢？我

们既无法在硬件的层次上进行比较，又无法在软件的层次上比较，就只好进入更高级的层次上来，比较依赖于软件的概念框架，也就是符号的网络。

研究大脑的思维活动要解决两个重要的问题。第一个问题要解释在低层次上神经细胞的激发怎样引起高层次上符号的行为。第二个问题就是如何创造一种理论能够解释高层次上引发符号的交流却并不涉及低层次上的神经细胞的活动。如果真能做到第二点，那就是为人工智能的研究提供了理论基础，即可以用不同于大脑的硬件来实现人的智能活动。

我们可以先考虑一下，我们怎样才能意识到计算机或大脑在低层次上的行为？在这样复杂的系统中是否有一种客观的方法从低层次的描述中抽出高层次的描述来。这一点对于计算机来说，答案是肯定的。但是对于大脑来讲又怎样呢？答案似乎也是肯定的。我们可以建立这样一种模型，用块的术语来粗略地描述大脑的运行状态。这是一种描述大脑状态和活动的语言。不过在一个人的大脑中潜在地存在着无限多的符号，也存在着无限多的引发这些符号的途径。外部环境则在选择哪一条途径时起着重要的作用。

这些符号可以彼此引发，不过要形成对于一个

事物的意识需要引发一组符号。这种意识是在比符号更高的层次上形成的，我们把它称为子系统的层次。所谓子系统是一组相互联系的符号的集合。虽然它和符号之间并没有明显的差别，与其他子系统之间也没有明确的界限。但是它确实自成一体。它一旦被引发，自身的机制就会起作用。子系统使用不同的符号，不同的子系统也可以使用共同的符号。这种共用的符号类似于计算机程序中的共用编码。因此这种符号的引发会产生什么样的结果，要根据它在哪一个子系统之中来定。

这种子系统还有一种重要的功能就是控制哪些符号被引发，以什么方式被引发。因此也可以说它是符号的符号，是表示符号活性的符号。

总之，作为思维的物质基础的大脑，它的活动表现出多层次的结构。而符号这一层次尤其值得我们注意。这也是与计算机科学、人工智能联系最密切的层次。

6.3 塔式基—丘奇—图林定理

我们已经看到，要理解大脑这样复杂的系统，只有靠组合成块的方法来构造较高级的层次。这样

做的结果是，每一层都要在精确性方面有所损失。最上面的那些层次可以说是“非形式系统”。这种系统遵循许多复杂的规则，我们现在还无法用语言来描述它们。这也是人工智能的研究想要揭示的内容。但是在这个系统的底层却是由简单的、可以形式化的规则所控制。

我现在要介绍一个重要的定理，它和数学、大脑和思维、人工智能等各个领域有着极为密切的联系。这就是著名的塔式基—丘奇—图林定理。这个定理描述了大脑系统最低层次的基本规律。

1936年美国的逻辑学家阿朗索·丘奇证明了这样的定理：

没有一种可靠无误的方法可以分辨数论中的真假陈述。

这就是说没有一种可靠无误的方法可以分辨形式数论系统中的定理和非定理。

塔式基定理则指出，对于数论的真理性不存在一种判定性的程序。这与丘奇定理是等价的。

图林研究计算机的基本理论也提出了与此等价的定理。

为了理解上述定理，有必要描述一下丘奇—图林命题。（通常称为丘奇命题）这涉及有关数学、

大脑和思维的最重要的哲学概念。

实际上，这个命题象茶一样可以具有不同的浓度，我们将给出它的不同形式并考虑它们的含义。

丘奇—图林命题，（重复形式）：

只有通过数学运算才能解决数学问题。

所谓数学问题是判定某些数是否具有给定的算术性质。所谓“通过数学运算”是指在确定一个数是否具有某种性质时只能采用少量的运算，如加、乘、检验相等还是不等，这些运算可以组合起来并重复地使用。这是我们探索数的世界时可以使用的唯一工具。

关于这个定理还可以表述成这样的形式：

丘奇—图林命题，（标准形式）：

如果有一种方法，并且有一种有智力的生物按照这个方法将数分成两类。再假设这种方法总能在有限的时间内给出回答，并且总能对给定的数作出同样的回答。那么一定存在有限的Floop程序（即某种一般递归函数）能给出与上述方法完全一样的回答。

这个命题并不是一个在数学定理的意义上可以证明的事实。它是一种关于大脑采用过程的假说。

为了区别个人和集体的思维过程，我们再给出一种比上述命题较弱的形式：

丘奇—图林命题，（集体过程的形式）：

如果有一种方法，并且有一种有智力的生物按照这种方法将数分成两类。再假设这种方法总能在有限的时间内给出回答，并且总能对给定的数作出同样的回答。附加条件：如果这种方法还能通过语言从一个有智力的生物传递给其他的生物。那么一定存在一种有限的Floop程序（即一般递归函数）能给出与上述方法完全一样的回答。

我们想在这儿介绍一位印度的数学奇才拉马努贾。他来自印度最南部的泰米尔纳德。拉马努贾没有受过高等教育，只在高中学过一点数学。有一天，有人发现了他的数学天才，就送给他一本有点过时的数学分析教科书，于是他就单枪匹马地冲击数学分析的领域。当时他只有23岁。有人向他介绍著名的英国数学家哈迪，拉马努贾就把自己最出色的成果写信告诉了哈迪。他根本不懂英文，连信封也是别人帮他写的。哈迪收到这框东西后的惊讶程度是难以形容的。他从未见过这样的东西，只要一看就会明白，只有第一流的数学家才能写出这样的东西

来。这肯定不是胡编的，没有人能够想象发明出这样的东西来。后来在哈迪的资助下，拉马努贾来到了英国，他们进行了紧张、认真的合作。可惜当他33岁时就因肺结核而夭折了。

拉马努贾有些奇特的地方不同于那些数学大师。他缺乏严密性，常常只是把结果简单地写出来。他一口咬定，这是出自一种模糊的直觉而不是严格的探索。他还常说在梦中得到了神的启示。这就使他带上了神秘的色彩。更神秘的则是有许多这类“直觉定理”实际上是错误的。有不少受过教育的人热忱地把拉马努贾的直觉能力作为一种证据，说明可以通过直觉的洞察去认识真理，他的出色能力也确实增强了这种信念。即使是否认拉马努贾具有神秘力量的哈迪也说过：“我不敢肯定，从某种意义上讲，他的失败不比他的任何成功更加奇妙。”

拉马努贾的另一个突出特点是“与整数的情谊。”当然有许多数学家也具有这种特点，但在拉马努贾身上表现得更加极端。哈迪记得，有一次他去探望病中的拉马努贾并对他讲，自己刚才乘坐的出租汽车号码1729似乎没有什么意义，但愿它不是一个不祥的预兆。拉马努贾却回答：“不，这是一个很有意思的数，这是可以用两种方式表示成两个

数立方和的最小数。”哈迪又问他，那么对于四次方来说，这个最小数又是什么呢。他想了想回答说，他还找不出很简单的例子，这个数一定很大，答案是：

$$635318657 = 134^4 + 133^4 = 158^4 + 59^4$$

但是，哈迪并不相信拉马努贾的数学才能有什么与众不同的秘密。他的观点可以表述成这样的命题：

丘奇—图林命题，（哈迪形式）：

归根结蒂，所有的数学家都是同构的。

我们还要给出这一命题的强化标准形式：

丘奇—图林命题，（同构形式）：

如果有一种方法，并有一种有智力的生物按照这种方法将数分成两类。再假设这种方法总能在有限的时间内给出回答，并且总能对给定的数作出同样的回答。那么一定存在某种有限的Floop程序（即一般递归函数）能给出与上述方法完全一样的回答。

而且这种思维过程与Floop程序在这种意义上讲是同构的。即在一定的层次上，在计算机与大脑中进行的步骤是相互对应的。

这并不是说大脑实际上执行了Floop程序。而是说大脑过程的顺序和Floop程序一样，它的逻辑结构可以映射到Floop程序里。

大脑是一种特殊的形式系统，在它的底层——神经细胞这一层——那些“规则”在起作用并改变它的状态，对于这些最基本的元件可能没有什么解释。而在它的上层出现了有意义的解释——将我们所谓的“符号”与现实世界对应起来。这有点象哥德尔结构，它的高层次同构具有一种高层次的意义可以读成一串。但是这种高层次的意义依赖于低层次——一旦引进哥德尔编码的概念，就可以从低层次导出。然而在大脑中，在神经细胞这一层次发生的过程并不依赖于实际世界的解释；它们并不模拟任何东西。它们仅仅是构成一种基础支持更高的层次。在这点上，它们更象是袖珍计算机中的晶体管。可以设想，在包括想象和类比在内的程序中，一定具有多层次的结构，而不是只有一种推理层次来把握概念的系统。

如果用计算机来模拟大脑的神经网络，我们就可以给出上述命题的微观形式：

丘奇—图林命题，（微观形式）：

生物组成部分的行为能够在计算机上进行模拟。这就是说，只要给出任何组成部分（最典型的可以假设是细胞）的内部状态与局部环境足够精确的描述，那么它们的

行为就可以用Floop程序来进行计算。

这个命题说明大脑的过程并没有什么神秘的地方。由此可以得出下面的引理：

丘奇—图林命题，（简化形式）：

所有的大脑过程都可以从可计算的基础上推导出来。

如果这个命题成立就可以为人工智能提供理论基础。当然，人工智能的目的并不是模拟神经网络，而是建立大脑较高级层次与人工智能程序之间的同构关系。而大脑的高级层次，即“符号”这一层次，又与宏观世界有着同构关系。

人工智能工作者至今仍然会提出这样的问题：“为了达到人工智能的目标，我们对大脑的复制究竟要细到什么程度？”这个问题的回答完全取决于你究竟要模拟人类意识的多少性质。

我们再来看一种唯灵论的上述命题：

丘奇—图林命题，（唯灵论的形式）：

有些大脑所能做的事，可以用计算机来模糊地逼近，但是大部分不行，其中最有趣的肯定不行。而且不管怎样，即使能够做到的也还有灵魂需要解释，而这是计算机绝对做不到的。

最后我们要给出这一命题的一种极端形式：

丘奇—图林命题：

计算机是可笑的。科学一般来讲也是这样。

持有这种观点的人并不少，他们不会欣赏在探索人类大脑的抽象结构中所具有的深度、复杂性和美。他们不相信能够解释大脑，更不相信计算机可能具有这种能力。因此，完全的理性主义者相信，任何大脑过程原则上都可以转换成计算机程序，但是其他的人可能会感到，象美这样的概念是很难定义并为计算机所接受。

不过，如果认为"非理性与计算机是不相容的"那就是一种误解。完全可以对计算机给定一些指令使它给出非逻辑的陈述。其实问题很简单。存在着不同层次的意义。神经细胞总是完成正确的运算并不能保证在它们基础上的高层次的结论也是正确的。我们没有理由认为计算机中有正确无误功能的硬件就无法支持高层次的符号行为，它们能表示疑惑、遗忘或者对美的欣赏这样一些复杂的状态。

这种层次的差别为我们提供了反击卢卡斯的新炮弹。卢卡斯认为哥德尔定理适用于可控的机器，因为作为机器最本质的东西就是它必须是一种形式系统的具体实例。

我们已经看到，这个论断在硬件的层次上来讲是对的，但是还有可能存在更高的层次，这就未必正确了。卢卡斯的话给人的印象是他所讨论的模拟思维的机器只有一种层次，它在这种层次上操作那些符号。

事实上人工智能的研究已经创造了许多“思维模型”，在它们的最高层次可以进行想象、描述类比、遗忘思想、用模糊概念进行操作等等。它们确实依赖于基础硬件的正确功能。从这种意义上讲，它们仍然是“形式系统的具体实例。”但是哥德尔定理对于它们是不适用的。

于是我们可以得出上述命题的一种最强的形式：

> 丘奇—图林命题，（人工智能形式）：任何一种思维过程可以用计算机程序来模拟，它的基本语言具有和Floop同样的能力——这就是说，所有的部分递归函数都是能够程序化的。

如果人工智能的研究者希望自己设计的程序具有人的思维能力。那么他们就应该使计算机在最低的层次上更加接近大脑的机制。

6.4 思 维

我们已经探索了进行思维活动的大脑机制。探索了这种机制的不同层次。那么哥德尔理论又为我们理解自己的思维提供了什么启示呢？在理解哥德尔定理证明的过程中，我们懂得了同构型的编码、解释意义的不同层次、自我反省的能力。这些都会帮助我们理解在思维过程中产生符号和对符号进行运算的过程。加深我们对于思维结构的不同层次之间联系的认识。

我们现在懂得了一个系统在高层次上所具有的解释能力，在低层次上是没有的。这就是说，有些事实只能在一个系统的高层次上来加以说明，而在低层次上却做不到这一点。

我们相信，要想比较合理地解释思维过程中所出现的一些现象，例如念头、希望、想象、类比、意识和自由意志，恐怕都离不开怪圈。这是思维结构的不同层次间的相互作用。上面的层次是靠底下的层次来支持的，但是又返回来影响和控制底层的活动。因此我们理解自己思维过程的关键在于了解我们大脑内部层次之间的自相缠绕。

就拿美感来说，形成美的感觉是一种重要的思维活动。现在的问题是，当我们欣赏一幅名画时，是怎样感觉到美的。难道是那些线条和点在我们视网膜上形成的形象？当然也可以这么说。问题是这种形象如何通过大脑的分析机制而产生美的感觉。因为我们所能感觉到的只是二维平面上的形象，然而引起我们感情反应的却是画中所表现的意义，这是内含在二维形象中的多维内容。要把这种多维的内容从中抽出来，显然需要某种复杂的机制，这种机制不同于可以形式化的判定机制。

一段音乐也具有双重含义。一方面它是四处弥漫的，和世界上的许多事物发生联系的，另一方面这段音乐又具有本身的含义。而要解决这个两难问题的关键在于我们如何考虑意义的翻译器，也就是说听众的心理机制。在演奏乐曲时，这种机制可以帮助听众从中抽出其含义来。当听众第一次听到这段乐曲时，这种机制就可以获得许多方面的重要含义。这似乎说明，这些含义是乐曲本身所包含的。但是这只是问题的一个方面。对于音乐含义的翻译是由多维的认识结构来完成的，这种结构可以寻找这段乐曲和其他多维心理结构的联系而把以前存在的信息也结合在一起。因为这些结构已经把过去的

经验进行了编码。随着这个过程的进行，它的全部含义才慢慢地被揭示出来。实际上，有的人要几年之后才懂得一段音乐的核心意义。这看来又支持那种观点，音乐的含义是四处弥漫的，靠翻译器把它们逐渐集聚起来。

还有一个很好的例子也可以说明意义的两重性。这就是用古代语言写成的碑文。一方面它的含义贮存在世界各国的图书馆和学者们的头脑中。另一方面它显然又是包含在碑文中的。

由此可见，在我们思维过程中进行分析的“形式”可以分成两种。一种是“符号型”的，另一种是“语义型”的。“符号型”的性质可以通过有限的试验过程来进行判断。“语义型”的性质却需要无穷无尽的试验。根据直觉，我们可以感觉到，“符号型”的含义是浮在形式的表面，因而并不涉及到多维的认识结构。而“语义型”的含义正好相反，它是深藏在形式内部的，需要靠多维的认识结构把它抽出来。要判定形式数论中的串是不是定理就是一种语义型的性质。因为我们无法运用某些标准来判定一个串是否为定理。要抽出一个串的内在含义实质上就是要确定它和所有其他串的相互关系。而做到这一点就需要无止境的试验。

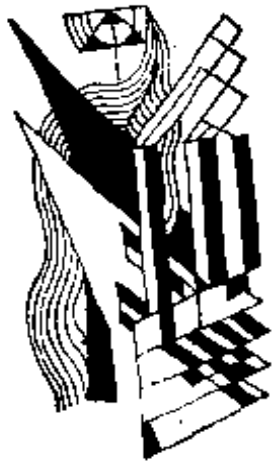
符号型性质与语义型性质的区别就在于符号型的性质是毫不含糊地存在于我们认识的对象之中，而语义型的性质依赖于它和无数其他对象的关系。符号型的性质原则上讲是确定的、没有什么秘藏的内容，秘藏的内容这个特点是属于语义型的性质。按照这种说法，美感显然不是一种符号型的性质。那么它是不是一种语义型的性质呢？我们先从一个人来考虑。几乎每个人都有过这样的体验，在某个时刻感到一样东西是美的，在另外一个时刻却感到它是丑的。这就是说，美是随着时间而变化的？当然也可以倒过来说，并不是美在变化，而是感觉的人在发生变化。那么能否说，对于特定的观察者、特定的画，在特定的时刻，美是一种确定的性质呢？或者这些说法本身就有问题呢？也许美感是因人因环境而异的，它可以在不同的层次上得到解释。这些不同的解释可以抽出不同的意义，建立不同的联系。由此可见，关于美的概念是很难确定下来的。同样，在数学中，真理性也是最让人头疼的概念。

最后，我要谈到关于真理的核心问题，即爱皮梅尼特悖论。在形式数论系统中重构这个悖论具有两个不同的层次。在第一个层次上，它是一个自我相关的句子，从它是对的可以推出它是错的，从它

是错的可以推出它是对的。在另一个层次上它又是一个有关整数性质的句子，它同样具有两难性。这种有关整数的陈述既不能是对的，又不能是错的。这就使我们联想到是否在普通语言下面也有一个层次。就是说有一个算术层次在支持着上层的含义。既然语言是以大脑机制为基础的，那么我们是否可以找到爱皮梅尼特悖论在神经细胞这一层次上的基础呢？这种想法告诉我们应该放弃这样的念头，大脑可以提供有关真理概念准确无误的表示。要提供有关真理的完整模型由于纯粹生理上的原因也是不可能的，因为它要求在大脑中所发生的事件都是不相容的。

7

人 工 智 能



7.1 回 顾

1950年艾伦·图林撰写了一篇有关人工智能的文章《计算机与智能》。这篇预言性、挑战性的文章一开始就提出了“机器能否思维？”这个问题。他主张用所谓“模拟遊戏”中的操作方法去探索这个问题。人们现在把这称作图林试验。图林在这篇文章中提出的许多光辉思想至今对于人工智能的研究还有深刻的影响。

人工智能的历史可以追溯到巴斯卡和莱布尼

兹，他们梦想有智能的机器。到19世纪，布尔和德·摩尔根提出了“思维定律”，这实质上是命题演算。这是走向人工智能的第一步。后来贝比奇设计了第一架“计算机器”，这是计算机硬件也是人工智能硬件的前身。人们可以把机器装置完成以前只能由人脑完成的工作算成人工智能的开始。我们很难体会当年人们初次见到转动的齿轮能够完成大数的加法和乘法时是怎样一种感情。也许他们会感到自己已经看到了“思维”在这些硬件中流动。不管怎样，我们知道，当一百年后电子计算机首次建造出来后，它的发明者确实感到有另外一种会思维的物体的存在了。至于真正的思维过程究竟是怎样的，至今仍然是一个谜。

有趣的是，仿佛有这样一条关于人工智能的“定理”：一旦某种思维的功能被编制成程序，人们就不再认为它是“实际思维”的基本组成部分了。而人工智能的核心总是指那些还未能编制成程序的部分。下面就让我们来回顾一下人工智能所研究的领域：

机器翻译

直接的（查词典、重排词序）

间接的（以某种内部语言为媒介）

游戏

象棋、桥牌等

证明数学定理

符号运算

图象处理

识别手写体

阅读各种字体的文章

听

理解口语

分出音节或词素

综合整个词或句子

理解自然语言

回答专业性问题

分析复杂句子

造句或写较长的段落

运用实际知识理解文章

分析意思含糊的关系

产生自然语言

抽象诗

任意的句子、段落或较长的文章

从知识的内部表达产生输出

思想性的创作或艺术创作

写诗

写小说

音乐创作

类比思维

几何形状

根据有关的领域作出某一数学领域的证明

学习

调节参数

形成概念

在人工智能研究开辟这些领域之前，人们无疑是把这些功能看成人类特有的智能。不管人们以后关于“真正思维”的概念发生什么变化，都改变不了这样的事实，人工智能的研究深化了我们对思维过程的认识，扩大了我们的思维功能。

机器翻译初看起来是件简单的工作，只要查阅字典重排词序就可以了。实际上却要复杂得多。这种困难不仅仅是缺乏有关成语之类的知识，而且涉及到有关领域的思维模式以及在这种模式中运算的符号。运用有关领域的模型进行翻译的方案也由于含义的模糊性和多义性而很快搁浅了。因此这仍然是悬而未决的问题。

计算机下棋也比人们最初的估计困难得多。已经弄清楚，人们在自己头脑中表示棋势并不是只要

记住每个棋子在什么位置以及下棋的那些规则就可以了。它还涉及到对几个棋子之间关系的理解以及关于启发式的知识，而这些是和更高层次的块有关的。

阿瑟·萨谬尔的下棋程序采用动态和静态两种方法来估价每一步的局势。静态方法是采用一个数学函数，有几个变量描述一种局势的特点，从而随时可以在实际上计算出来。而动态方法则是构造一棵“树”，即由每一步后可能走的步构成的树。萨谬尔的程序就是依据这些方法的估价来确定下棋的战略，决定走哪一步棋。这两种方法又是相互依存的，每一种方法的改进都使另一种方法受惠。

那么能否用这种程序来下象棋呢？根据实践得出的结论是，用这种程序下象棋的困难要比人们最初的设想大得多。

说到计算机创作音乐，究竟谁是真正的创作者呢？是计算机还是人？这是人的智能，计算机不过是创作的工具罢了。

回顾人工智能的早期历史，人们最早作出的努力之一便是企图实现定理的机器证明。将问题进行归纳是其中一项重要的技术。也可以将整体的问题转化为局部的问题。然而这种方法并非总能保证成

功的。如果我们把对问题的探索看成是在问题空间中逼近确定的目标，那么这就是一种抽象的概念性空间。当你认识到如果直接奔向目标会遇到抽象的“篱笆”时，你可以采取两种行动：（1）你以某种随机的方式离开目标，希望能够发现一道隐藏的“门”；（2）试图去发现能够表示问题的新“空间”，在这新空间中在你和目标之间没有抽象的篱笆隔开。这虽然是一种理想的方法，但如何选择一个好的空间来表达问题却是一个非常复杂的问题。

数学也是人工智能可以研究的一个有趣领域。每个数学家都有一种感觉，在数学的各种思想之间可以有某种度量。这就是说数学的成果构成了一张巨大的网，而两者之间往往有许多环节。有时数学中的两个定理很接近，给出其中一个很容易证明另一个，或者两者很相似，甚至是同构的。但这只是一种主观的感觉，能否有一种客观的标准呢？有时在不同分支中的两个定理似乎是不相干的，但是以后的进展改变了人们的这种看法。如果我们能够把对数学中紧密程度的感觉编制成程序，那么也许我们可以创造出一位原始的“人工数学家”来。

人工智能有着广泛的应用，而它的核心问题就是如何表达知识。在人工智能研究的早期，人们假设

知识是以一条一条句子贮存起来的，而要把知识以一条一条地移到程序中去的最好方法就是发展一种简单的方法把事实翻译成小小的被动的数据。每个事实是一条数据，当程序用到它时能够提取出来。象棋程序就是一个例子，它把棋势进行编码，然后有效地贮存在“记忆”中，需要时用子程序来调用。

然而心理学家们早就发现，人类是以一种更为复杂的方式来贮存事实的。最近人工智能研究者又重新证实了这一点。他们现在面临着如何使知识“块化”的问题，和程序型知识与陈述型知识之间差异的问题。这些问题又与在反省中能够理解与不能理解的知识之间的差异有关。

其实，假设所有的知识编码成被动的数据是与计算机设计最基本的事实相矛盾的。如何进行加法、减法、乘法等等运算并不是编码成数据贮存在记忆中的，而是表现在硬件的线路连接方式上。

重要的是在原则上说明这一点。不管是以数据的方式还是以程序的方式来处理信息，只要我们不考虑处理的效率，那么用一种方式可以做到的，用另一种方式也能做到。但是理智似乎在告诉我们，一种方式比另一种方式更为优越。

也许我们可以从自然界的生物获得一点启示。

为什么在蛋白质中表现出来的遗传信息要在DNA中编码呢？答案是：将同样的信息以几种不同的形式用于不同的目的是非常有用的。将遗传信息以模具和数据的形式贮存在DNA中有一个优越性就是很容易将两个个体的基因组合成一个新的基因型。如果信息只是贮存在蛋白质中是很难做到这一点的。把信息贮存在DNA中的另一个原因就是这样很容易转录或转译到蛋白质里。当它不需要时并不占据很多地方，一旦需要时就可以作为模具使用。

如何以最好的方式进行知识的表达和运算，关于这种思维活动的研究有许多学派。其中很有影响的一派认为是用形式化的概念来表达，当然这种表达的基本运算也就采用形式化的演绎推理。

但是人还有一种能力就是能对类似的情况进行比较，我们可以把这称为类比的认识。这在人的智力中占有重要的地位。

当然，人类的记忆也是极为复杂的。有的知识经过编码后贮存在复杂的系统内，却可能没有一种程序，一种规则或者其他提取数据的方法可以使你把它提取出来。这时我们就会说它被遗忘了。因为提取它的方法暂时或者永远失效了。但是所谓遗忘，往往是在较高的层中失去了线索，并不意味着所有

的信息都丧失了。这些信息完全可能在较低的层次上依然存在。

我们再来看看计算机程序是怎样做到某种程度的“理解”。

7.2 人工设计的游戏

托里·威诺格拉特（化名托尼·厄尔威格）在1968—70年设计了SHRDLU程序。当时他还是麻省理工学院的博士研究生，正在研究语言如何和理解相结合的人工智能问题。为了说明这种程序的功能，我们来看这样一段经历。

有一天埃塔·奥茵偶然撞进了麻省理工学院的人工智能实验室。那里的计算机刚刚采用新设计的SHRDLU程序。于是在奥茵和计算机之间进行了一次有趣的对话。而程序的设计者托尼·厄尔威格博士则对他们的对话作了必要的注解。

1.奥：拣起一块大的红方木块（参阅图34）

机：好。

博：计算机执行命令就回答一声“好”。要执行命令就得把上面那块绿的移开。

2.奥：抓住那块锥体。

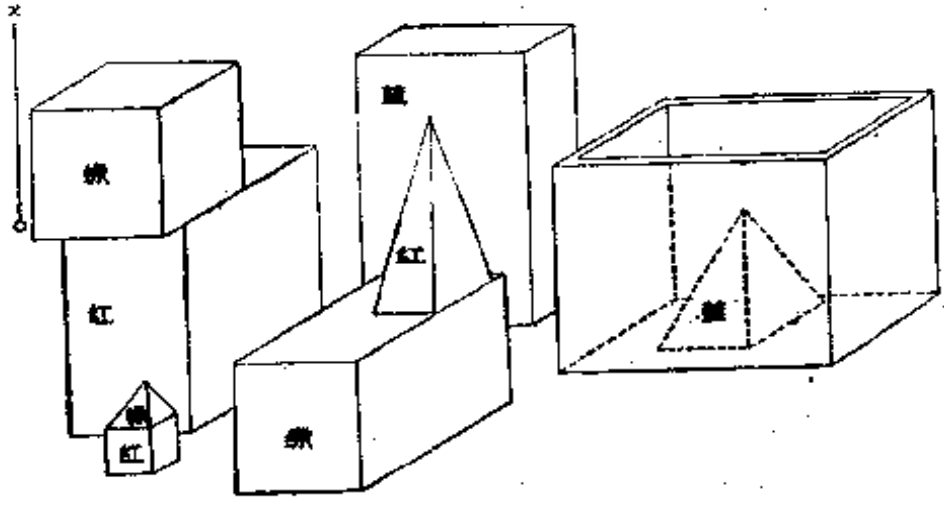


图34

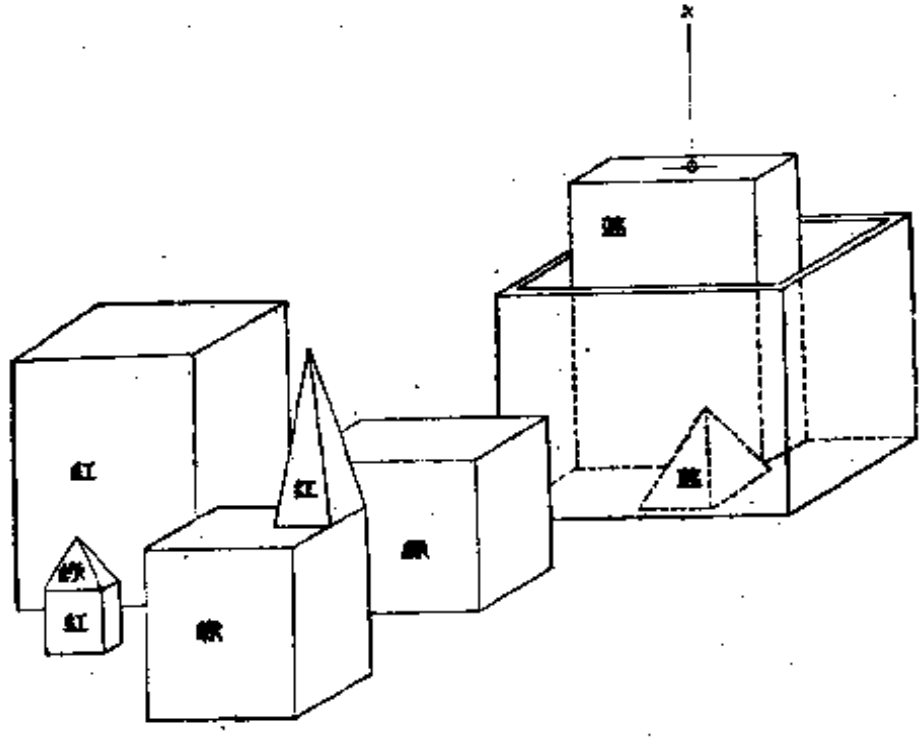


图35

机：我不明白你指哪一个锥体。

博：可以看到这里共有3个锥体，它不知道奥茵指的是哪一块。

3.奥：找出比你拿着的那块更高的方木块，把它放进盒子里去。（参阅图35）

博：计算机必须先弄懂“一块”和“它”是什么意思，然后才能去推论应该拿哪一块。

机：我设想你所指的比我拿着那块更高的木块是指“它”。

博：这里的“它”本来可以指挑出来的那块，也可以指拿着的那块。计算机通过试探的方法在两者之间进行选择。它必须把红木块放下，拣起更高的蓝木块。

4.奥：盒子里有什么？

机：蓝锥体和蓝方木块。

5.奥：锥体底下是什么？

博：计算机不再问前面提出过的问题，而设想你所指的就是刚才提到的那一块。

机：盒子。

6.奥：在盒子外有多少方木块？

机：4块。

7.奥：其中至少有一块比我刚才让你捡起来的那块

宽吗？

博：象“其中至少有一块”这类词组也应该是指木块。而记忆可以帮助计算机推出“刚才让你捡起的那块”指的是什么。注意这个词组的复杂结构。而SHRDLU的语义分析器所采用的办法是按照一般的方式去把握这种复杂性的。

机：有，是那块红的。

8.奥：红方木块底下是什么？

机：桌子。

9.奥：桌子能捡起木块吗？

博：SHRDLU的语义学规则应该包括“捡起”的知识。只有动物才能做到这一点，而桌子并不是动物。

机：不能。

10.奥：方木块能支撑锥体吗？

机：能。

博：SHRDLU的推理系统发现了这样的实例，因而知道这是可能的。

11.奥：锥体能支撑锥体吗？

机：不知道。

博：计算机找不到这样的实例，也无法在程序中找到，虽然它最终能够改进这种能力。

12奥：把两个锥体叠起来。

机：我做不到。

博：它作了尝试，但是失败了。

13奥：那个蓝锥体是属于我的。

机：我明白了。

博：所有关于颜色、形状、大小和位置的信息，计算机原来都知道。但是“属于”这种性质原来是不知道的。现在才贮存进去。

14奥：所有不是红色的方木块都属于我，但锥体下的东西不属于我。

机：我明白了。

15奥：盒子属于我吗？

机：不。

16奥：盒子里有什么东西属于我吗？

机：蓝色方木块和蓝色锥体。

17奥：请你堆起两块红色方木块，另外再加上一个绿色的立方体或者一个锥体。

机：好的。

博：这些逻辑连词“和”、“或”等，既要从语法上去把握又要从语义上去把握。

18奥：把最小的锥体放到顶上去。

博：在它的词典里并没有“最小”这个词，但是

可以通过小米理解。计算机在比较它们的大小之后完成了这一步。

机：好的。

19奥：在绿色立方体顶上有多少块？

机：我不敢肯定“在顶上”指的是什么。

博：这和放到顶上不同，因为放到顶上只有一种可能性，否则都做不到。“在顶上”也许指的是正前方，也许指的是上面的任何地方。

20奥：做一个尖塔。

机：我不懂“尖塔”这个词。

博：当计算机碰到一个它还不懂的词时，立刻就中断了。

21奥：两个绿色立方体和一个锥体组成一个“尖塔”。

机：我明白了。

22奥：做一个。

机：好的。

23奥：谢谢你。

机：欢迎你以后再来。

通过这段人—机的对话，我们可以清楚地看到，经过十多年的发展人工智能已经到了什么

水平。机器变得多么聪明。它会运用已有的知识进行推理，也会接受新的概念，学习新的知识。

7.3 概念框架与模式识别

人们的思维中常常有稍稍改变某种条件的倾向，并设想这种变动会导致的后果。

如果我们借用数学中的术语：常量、参量和变量，就可以用来描述稳定性的不同层次。一个事件可以看成是发生在某个关系的网络中。最底层的是相互关系最深的方面，有时是完全不能变的。例如，我们中的大多数人从未想到过要改变我们世界的三维空间结构。还有一些层次是暂时建立起来的，并不是永久性的。它们固定了某些条件。我们不妨称之为基本假设。这是我们思维的基础。虽然我们知道这是可以改变的，但是一般来讲，我们宁可把它们看成是不变的。这些我们仍然可以称之为常量。例如一场球赛的规则就是这类常量。还有一些“参量”可以认为更容易变化，但是暂时让它们保持不变。对于一场球赛来讲，参量包括天气、竞赛的球队等等。参量的层次可能有好几层。最后则是最容易变化的变量。

在人工智能中要用到框架的概念。框架里面还套有子框架。这样就可以按照你所需要的精细程度选择合适的框架。这种套在一起的框架结构就象是一组关于美国公路的地图册。最上面是全国的地图，里面则是分州的地图。如果你要更详细的，甚至还有某些较大城市的市区地图。你还可以设想任意精细的地图册，乃至于一个街道、一幢楼房、一间屋子的平面图。也就是说，象是通过具有焦距可变镜头的望远镜来观察，其中每个镜头都有各自的用处。每一个框架又有自己的可变层次。因此，当我们把一个框架塞到另一个框架里去时，就有可能产生冲突。作为例子我们来看看所谓的“蓬加特问题”。

“蓬加特问题”就是模式识别问题。解决这种问题的程序可以分为几步，它们逐渐将原始的数据变成描述。最初的几步是固定的，以后的几步就逐渐变得灵活了。最后一步则具有试验的性质了。

对原始的数据进行预处理就意味着测出它的一些突出性质。而这些性质的名称就构成了这一问题的“最小词汇表”。主要性质词汇表的典型术语如下：

线段、曲线、水平、垂直、黑、白、大、小、

尖、圆……

而在预加工的第二阶段就使用了有关基本形状的知识，这种术语有：

三角形、圆、正方形、锯齿形、隆起状、直角、顶点、岐点、箭头……

试验性的描述则可以用这样一些术语：

在上面、在下面、在右面、在左面，在里面、在外面、接近于、远离、平行于、垂直于、在一条线上、分散的、平均分布、不规则分布的等等。

值得指出的是，这种描述虽然可以使一个人识别某种图形，但是并不能使他将原图重构出来。这有点象音乐的风格。你可以确定无误地认出莫扎特的作品，但是却无法写出任何东西可以欺骗别人，使人认为这是莫扎特的作品。

有一种很好的策略就是尝试描述彼此在结构上的相似性。从某种意义上讲，这是有可能的。这里有两点很重要。第一是关于“描述图式”的思想，或者说是图样；另外一点就是关于“同样性检验器”的思想。

检验同样性可以在一个描述层次中进行，也可以同时在多种描述的层次中进行。而描述图式或图样最重要的是选择一种方法，将有关的描述分解成

各部分，使它们能反映出所有图中的共同性。

要解决“蓬加特”问题还需要概念的网络，在这样的网络中我们可以看到，

“高”和“低”是相反的。

“上”和“下”是相反的。

“高”和“上”是类似的。

“低”和“下”是类似的。

“左—右”的差别类似于“高—低”的差别。

“相反”与“类似”是相反的。

此外我们还可以看到，

正方形是多边形。

多边形是闭曲线。

闭曲线有内部和外部。

“内部”与“外部”是相反的。

概念网络的一个主要作用就是使最初的一些错误概念能逐步得到改进。

在识别过程中，另一个重要的部分就是抽象的层次与元描述。这就是说存在着比几何描述层次更高的抽象。这就是关于描述的描述。

现在让我们来考虑另一个问题。究竟怎样去掉一部分信息呢？这里涉及到两个概念，我称之为“聚焦”和“滤波”。所谓“聚焦”就是集中描述图中

的某些部分而把其他部分抛弃掉。所谓“滤波”是采用某种特定的观察方法看待图中的内容，而置其他方面不顾。这样它们就起到互补的作用，聚焦与观察对象有关，而滤波则和概念有关。

为了识别图象就需要有概念网络。概念网络则和符号系统有关。这就涉及到符号的创造、操作和比较的问题。

说到创造就不是一种机械的活动。有一种普遍的观念认为，随意性是创造的必要组成部分。事实上，偶然性确实是思维固有的特征。

从对“蓬加特”问题的研究，我们可以看到，人类通过遗传获得的对于模式的难以捉摸的理解包括表达知识的所有机制，如交错的相互关系，概念框架、概念的映射、可以变动的描述和元描述以及它们之间的相互作用，符号的融合和裂解、多重的表达、有欠缺的预测等等。

总之，人类的智能是如此缺乏规则。许多人工智能的专家风趣地说，他们自己也不明白当初怎么就糊里糊涂地闯到这个领域里来了。人工智能确实是一个大大有待于进一步开发的领域。

8

自我相关与自我复制

8.1 遗传机制的形式化

哥德尔定理的证明离不开自我相关。如果我们把产生自我相关的机制与生物体进行自我复制的遗传机制进行一番比较，就会发现两种机制之间有着优美的同构关系。

我们先来看比较熟悉的自我相关的结构，以下面这些句子为例。

（1）这个句子中有9个字。

（2）这个句子是没有意义的，因为它自我相

关。

（3）这个句子中没有副词。

（4）这个句子是错的。（爱皮梅尼特悖论）

（5）我现在写的这个句子是你现在读的。

这些句子的自我相关是很明显的。每个句子都在描述本身的性质。做到自我相关最一般的方法就是前面提到的奎因法。它是通过描述句子的另一部分来实现自我相关。而被描述的部分与奎因句子本身又是同构的。

我们知道，有一种计算机程序可以设计有同样能力的计算机程序。我们可以说这是一种自我复制。当然最典型的自我复制要数生物体的自我复制了。这种自我复制的机制并不是产生和原来一样的生物个体。因为在这种复制中保留的并不是某一个生物体的信息，而是同一物种的信息。

生物体的自我复制或者说生物的遗传机制是20世纪最迷人、最深刻的研究课题之一。我们并不想在这儿讨论有关遗传的生物学研究成果，而是要建立一种遗传模型。即用一套形式符号及其运算规则来描述遗传的机制，并把这种机制与自我相关的机制进行比较。我们用4种符号来表示遗传基因中的4种碱基：

A.腺嘌呤

G.鸟嘌呤

C.胞嘧啶

T.胸腺嘧啶

由这些符号组成的序列称为遗传模型中的串。

这些串可以进行各种运算，包括切断，接上或者保持原有的长度。而能够执行这些运算的机制我们称为酶。酶有一个显著的特点就是它对4种碱基中的某一个有所偏爱。

还有一个新的术语是复制模式。这种模式不是用A复制A，而是用A复制T或者相反，用G复制C或者相反。我们把这两对称为互补的碱基对。因此所谓复制一个串并不是复制它的原样，而是复制与它互补的串。

酶所执行的指令可以分成15种：

cut——切断串

del——从串上去掉碱基

swi——把酶转换到其他串上

mvr——向右移动一个单元

mvl——向左移动一个单元

cop——开始复制

off——停止复制

ina——把A插入这个单元的右边

inc——把C插入这个单元的右边

ing——把G插入这个单元的右边

int——把T插入这个单元的右边

rpy——向右寻找最近的嘧啶（T或C）

rpu——向右寻找最近的嘌呤（A或G）

lpy——向左寻找最近的嘧啶

lpu——向左寻找最近的嘌呤

我们可以把这些指令看成是氨基酸。每一种酶都是由这些氨基酸组成的序列。

例如这样的酶：

rpu-inc-cop-mvr-mvl-swi-lpu-int

作用于以下的串：

TAGATCCAGTCCATCGA

假设这种酶偏爱G，我们从中间的那个G开始。向右找到A在它的右边插入C，然后进行复制，右移，左移，然后转换到另一段串上得到：

GV

TAGATCCAGTCCACTCGA

为了清楚起见可以把它们例置一下，继续以下的程序得到最后的两个串：

ATG和TAGATCCAGTCCACATCGA

（注意：在进行int操作时，除插入T外还要在另一个串的相应位置插入互补的A）。

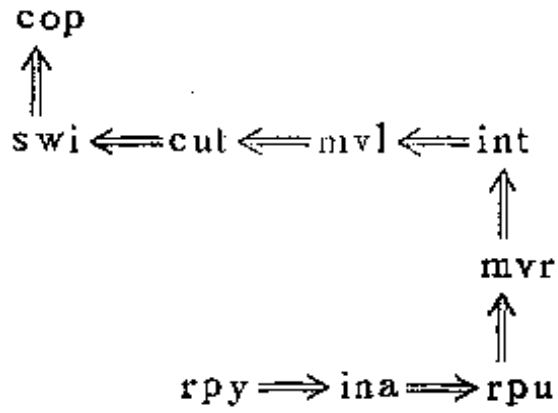
从这一例子我们可以看到，酶是怎样作用于那些串的。这就好象是MU系统中的规则，按照这些规则就可以产生新的定理。所不同的是在酶的作用下，原来的串永远消失了。

我们进一步确定这些氨基酸与碱基对的对应关系。这种关系可以列表如下：

	A	C	G	T
A		cut_s	del_s	swi_r
C	mvr_s	mvl_s	cop_r	off_l
G	ina_s	inc_r	ing_r	int_l
T	rpy_r	rpu_l	fpy_l	lpu_l

按照这种关系，上面例子中的酶就可以从最后得到的长串“转译”过来。

在表中每一格的右下角还有一个字母。这些字母以一种特殊的方式决定了酶的第三层次的结构。s表示保持原来方向，r表示向右转l则表示向左转。还是以我们刚才的酶为例子，可以得到如下所示的图：



根据这种结构中最初和最后的箭头方向则可确定这种酶偏爱哪一种碱基。

原始箭头	最终箭头	结合的字母
⇨	⇨	A
⇨	⇧	C
⇨	⇩	G
⇨	⇦	T

由此可见这种酶是偏爱C的。

不过上表中在AA这一格是空白的，它起标点符号的作用，表明一种酶的结束。这就是说，在一个串上可以记录几种酶。

到这儿为止，我们已经给出了遗传模型的所有规则。用这些规则去做一些符号游戏是非常有趣的，特别是那种可以自我复制的串。这就是说给定一个串，在核糖体的作用下就产生记录在串上的酶，这些酶作用于原来的串就生成子串。子串通过

核糖体的作用又生成第二代的酶，它作用于子串生成第三代的串。如此循环往复无穷。

因此这种遗传模型的核心是酶和串的相互变换。串通过核糖体译成酶，酶通过符号运算生成串。这样串就同时起了数据和作用这些数据的程序的作用。于是两种不同的层次就混在一起了。相比之下，在MIU系统中，推理规则在比定理更高的层次中，两者是分开的。TNT和其他形式系统也是这样。但是在TNT中不同层次在另一种意义上又混起来了，使语言和元语言的差别消除了。这就是关于系统的陈述可以在系统内部反映出来。

当然，实际的遗传机制要复杂得多，但是上述的遗传模型的“概念框架”有助于读者理解真实遗传的秘密。

DNA的原意是脱氧核糖核酸，它由一串比较简单的组分即核苷酸组成。每个核苷酸分子由三部分组成：（1）磷酸基。（2）去掉一个氧原子的核糖，（3）碱基。从碱基上可以将四种核苷酸区分开来。这四种碱基便是：A，腺嘌呤；G，鸟嘌呤；C，胞嘧啶；T，胸腺嘧啶。DNA通常都是双股的结构，两股上相应的碱基是互补的。在同一股上的核苷酸以共价键紧紧相连而形成骨架。而两股

相应的核苷酸之间则是较弱的氢键。

DNA在细胞核中，但是细胞中的许多“动因”是在细胞质中，特别是那些酶，它们是由核糖体在细胞质中制造的。但是关于酶的结构的信息却贮存在DNA中。这种信息是通过信使核糖核酸mRNA传给细胞质中的核糖体，使它完成创造酶的工作。

酶是蛋白质的一种。蛋白蛋由一系列的氨基酸组成。氨基酸有20种，这和前面的符号稍有出入。氨基酸的复杂程度与核苷酸差不多。但是蛋白质的序列则要短得多。典型的蛋白质分子由三百来个氨基酸组成，而DNA的串则由几十万甚至几百万个核苷酸组成。

DNA的双股解开并复制mRNA的过程为转录。mRNA在细胞质中遇到了核糖体就发生转译的过程。如果我们把mRNA比作一条磁带，那么核糖体就是录音机的磁头。所奏出的音乐就是根据上面信息所生成的氨基酸构成的蛋白质。我们把这些信息称为遗传密码。

由核糖体生成的蛋白质可以越来越长。它们不仅具有氨基酸序列的一级结构，而且具有相缠绕的三级结构。正如前面已经指出的那样，三级结构是蕴涵在一级结构里的。

可以把蛋白质比作音乐。音乐并不只是一串音符。我们可以从更高的层次上感受到音乐的旋律和速率。同样，蛋白质也有不同的结构层次。除了一级结构和三级结构外，还存在着二级结构和四级结构。不过我们在这儿就不详细谈了。

8.2 遗传机制的支持系统

生物界在漫长的进化过程中发展了遗传机制。这是大自然最伟大的杰作之一。我们在上一节中建立了使遗传机制形式化的遗传模型。这使我们能够从一个全新的角度来认识生物遗传的过程，从而可以获得许多有益的启示。现在让我们来考虑这样一个问题，“究竟什么样的DNA能够指导自己的复制？”当然并不是任何DNA都能够做到这一点的。DNA能够进行自我复制就必须包含一组指令，能够精确地把这样一些酶组装起来，而这些酶则能够完成这种任务。因而不能指望DNA的串能单独复制自己；要想从DNA中产生这些潜在的蛋白质不仅需要核糖体，而且需要DNA聚合酶。这就是说需要有一个系统来支持转录和转译的过程顺利地进行。我们把能够最低限度满足这种需要的系统称

为“最小支持系统”。对于DNA的自我复制来讲，这样的“最小支持系统”必须包括：

（1）某些RNA的聚合酶，它们能从DNA制备mRNA；

（2）某些核糖体。

我们不难看出，所谓“足够强的支持系统”与“足够强的形式系统”在本质上是一样的。只是两种不同的表现形式而已。前者是产生自我复制的先决条件，后者是实现自我相关的先决条件。这两种系统之间存在着同构关系。

我们不妨把遗传模型中的一些概念和形式数论系统中的相应概念作一番比较：

DNA的串	⟺TNT的串
mDNA的串	⟺关于数的陈述
蛋白质	⟺关于元TNT的陈述
作用于蛋白质的蛋白质	⟺关于元TNT陈述的陈述
作用于（作用于蛋白质的蛋白质）	关于（关于元TNT陈述⟺的陈述）的陈述
转录（DNA⇒RNA）	⟺ 翻译（TNT⇒N）
转译（RNA⇒蛋白质）	⟺ 算术化（N⇒元TNT）
遗传编码	⟺哥德尔编码

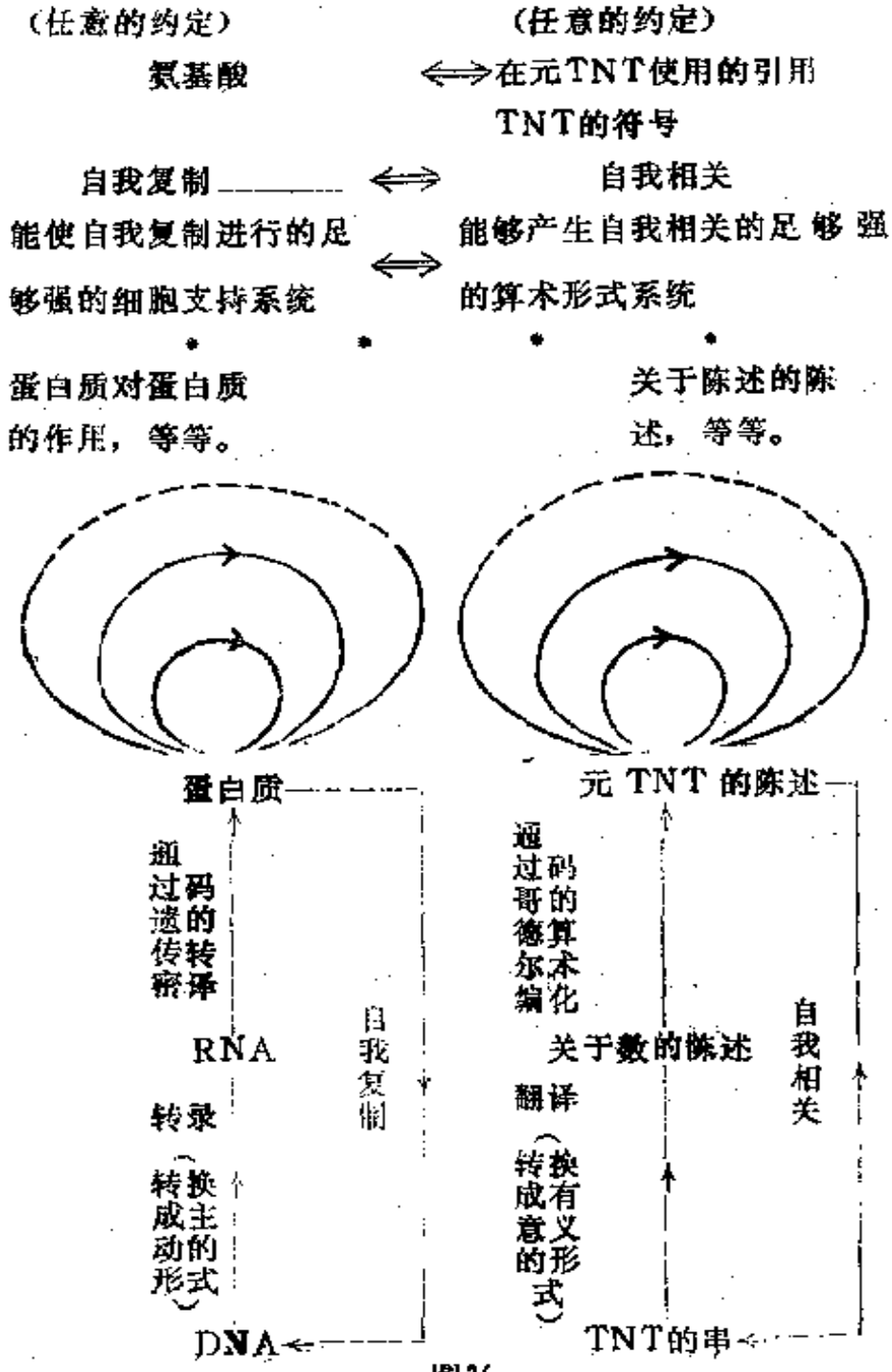

（任意的约定）
氨基酸
（任意的约定）
⟺在元TNT使用的引用
TNT的符号
自我复制
⟺
自我相关
能使自我复制进行的足够强的细胞支持系统
⟺
能够产生自我相关的足够强的算术形式系统
蛋白质对蛋白质的作用，等等。
关于陈述的陈述，等等。
蛋白质
元 TNT 的陈述
通过遗传密码的转译
RNA
转录（换主的式转成动形）
自我复制
DNA
通过哥德尔编码的算术化
关于数的陈述
翻译（换有义形式转成意的）
自我相关
TNT的串

图36

我们可以看到，在这两种抽象结构之间有一种几乎是神秘的共同性。当然这种对照并不是两种理论一致性的严格证明，但是它明显地表示两者之间有一种极为密切的关系，它值得我们进一步加以探讨。

从上面的图中可以看到，两边都出现了任意复杂的圈，一边是蛋白质对蛋白质的作用可以无限地延续下去，另一边则关于陈述的陈述也可以无限地延续下去。这就有可能在较高的层次上出现怪圈并且循环下去。

我们在第2章给出了唱机唱片系统与形式系统之间的对应关系。在这里又给出了遗传机制与形式系统之间的对应关系。有兴趣的读者可以根据这两种对应关系去建立唱机唱片系统与遗传机制之间的对应关系。

当然在遗传机制中还有很多复杂的现象，如病毒的自我装配、细胞的分化、繁殖的方式等等。这些概念也和转录、转译一样有助于我们对极其复杂系统的理解。例如细胞一再复制的过程引起了大量有专门功能的细胞的形成。这就好象字母的序列，从一个人传递到另一个人，尽管要求每个人都不使它走样，然而却总会使它带上传递者个人的影响。我

们还可以进一步想到，不但程序和数据是交织在一起的，而且程序的翻译机和数据的处理机，包括语言在内，也都密切地结合在一起。虽然可以在它们之间刻出界线、分出层次来，但是更重要的是要看到它们之间的交互作用。

我们可以用计算机的语言来描述分子生物学中的许多概念。因为DNA中包含了有关蛋白质构造的全部信息，而这些蛋白质是细胞中的主动因素。我们可以把DNA看成是用高层次语言写成的程序。它可以翻译成细胞的“机器语言”即蛋白质。另一方面 DNA 本身又是被动的分子，要受各种酶的控制。从这种意义上讲 DAN 完全象是一个数据。而且 DNA 还含有模板，可以复制出 tRNA（转移核糖核酸）这就意味着DNA还含有比本身更高层次的语言。

再拿蛋白质来讲。蛋白质是有主动性的分子，可以执行细胞的所有功能，可以把它们看成是细胞（细胞本身是处理机）的“机器语言”的程序。但从另一方面来讲，蛋白质是一种硬件，而程序却是软件，因此把蛋白质看成处理机也许更为合适。此外蛋白质往往受到其他蛋白质的作用，因此也可以把它看成是被处理的数据。最后我们还可以把蛋白

质看成翻译器。

我们也可以这样来分析核糖体和转移核糖核酸等。不过对于所有这些生化分子相互关系的分析仍然是很浅显的。我们只是想说明，我们认为截然不同的层次往往是交织在一起的。实际上计算机科学早就有一种明显的趋势把这些看来截然分开的各方面结合在一起。对于人工智能的研究来讲，这点就更为重要了。

8.3 没有结束的唱机之战

（阿基里斯去拜访蟹的家。蟹向阿基里斯展示了自己最崇拜的画家马格里特的作品。然后又向阿基里斯谈起乌龟与他之间的那场唱机之战。

蟹：你知道乌龟有多鬼，他几乎能破坏我所能得到的任何一种唱机

阿：不过据我所知这场竞争的结局是你最后拥有一台不可战胜的唱机——它自身带有摄象机和微电子计算机等等，从而能够将自己拆散并重组，使得自己不会被破坏。

蟹：可惜我的计划成了泡影。乌龟还是胜我一筹，

他看出了我忽略的一点：控制拆散和重装过程的部件必须在整个过程中保持稳定。因为显然的原因，它是无法将自身拆散或重装的。结果乌龟制造了一张唱片使这部分被震坏了。如果你不见怪的话，可以说乌龟总能发现一个系统中的“阿基里斯脚踵”。

阿：你一定感到很悲伤吧。

蟹：有一阵子是那样，不过幸好事情没有就此完结。在乌龟的推荐下，我借阅了一本包罗万象的书。这本书中谈到了能够组装自身的核糖体。也就是说，在一个细胞里有这样的生物单元，它们可以分裂并自动地组装起来。

阿：这可真妙！是否一架唱机也能有这种性质呢？如果这样你不是就可以有一架摧毁不了的唱机？即使被破坏了也能再自行组装起来。

蟹：我也是这么想的，并把这种想法告诉了唱机制造商。几个月之后，他写信给我，告诉我这种唱机终于造出来了。我满怀信心地邀请乌龟来试验一下我这架唱机。

阿：我想你这次总该如愿了吧。

蟹：很遗憾，当那架唱机的弹簧和导线满天飞开的时候，我是多么悲伤。乌龟为了安慰我就向我

解释，对此他也无能为力，这就是因为有了哥德尔定理。

阿：我实在很同情你，不过后来你怎么办呢？

蟹：我放弃了那种“完备唱机”的念头。我想一架唱机要不被毁坏就只好放那些不会损坏它的唱片，也就是说只演奏一些唱片。

阿：这就是说，你为了保存唱机就牺牲了那种想复制一切声音的愿望。

蟹：我是被迫这样做的。这就是防止唱机去演奏那些不相容的唱片。

阿：这可真是一个出色的战略。

蟹：不过乌龟也改变了他的策略。他捉摸怎样使唱片能逃避我的检查。

阿：那你怎样把那些“不相容”的唱片排除掉呢？

蟹：如果你替我保守秘密的话，我就告诉你。我的办法是把我所有的唱片秘密列成表。然后用电视机扫瞄这张表，用电子计算机记录这些数据。

阿：我想你的这种方法很容易失败。乌龟只须把你这张表弄到手就可以了。

蟹：没那么容易，因为我的摄象机会检查每张唱片的纹沟，然后交给计算机分析，只有具有我自己

风格的唱片才能通过这种检验。而这种风格是乌龟无法模仿的。你是否愿意看看我的这种防御装置的内部是怎样工作的呢？

阿：那当然好。我只要把镜头对准哪儿，屏幕上就会出现那些视界里的东西。

蟹：要是你把它对准电视机的屏幕呢？

阿！啊，我看到了一连串的屏幕，这可真是把自身卷了进去。而屏幕成了自我缠绕系统中的一部分。不过我的头被你弄成了一团乱麻。

蟹：我很理解你的感受。你为什么不抛弃那些关于自我缠绕的念头，坐下来看看我的藏画。

（蟹取下嘴里的烟斗，把它放在马格里特写有一行字的画上。）（图20）

阿：这张画真有意思。我喜欢它的框架，特别是框上的雕花。不过那一行字是什么意思？

蟹：这是一行法文，意思是："这不是烟斗"。

阿：可这确实是烟斗啊。你刚才还用它抽烟呢！

蟹：噢，你误解了。法文的"这"是指画，而不是指那个烟斗。当然烟斗还是烟斗。

阿：我可不知道画上的"这"是指整个画还是指画上的烟斗。我的天！这又是一种"自我缠绕"！我真感到不好受，我想我是要病了……

8.4 自相缠绕的层次

思维是最复杂的现象之一。人们在思维过程中往往改变自己的思维规律，甚至改变那些使这些规律变化的规律。但是这些都是指软件的规律。在最低层次的规律是永远不变的。神经元总是按照同样简单的方式进行操作。正是大脑中硬件的这种固定性造成了软件的灵活性，这就是智能的有趣结构。我们现在要讨论的就是这种可以自我改进的软件与固定不变的硬件之间的差异。

我们仍然以下棋为例。一般来讲，下棋的规则是不能变的，改变的是每一步棋之后的局势。让我们想想能否把下棋的规则也变动一下呢？能否重新定义马的走法，或者使棋盘增添一些方格？如果能够做到这一点，那么我们就有两种层次的规则。一种层次的规则告诉我们如何移动各种棋子，另一种层次的规则告诉我们如何改变这些规则。如果我们走得更远，用两种方式来解释各步棋：（1）作为下棋的步骤；（2）作为移动棋子的规则。那么当我们下棋时，每当我们移动一步棋也同时改变了下棋的规则。这样一来，下棋的规则总是在不断的变化

之中。于是就出现了自相缠绕的层次，或者说形成了怪圈。每走一步都要影响到下棋的规则，这些规则的变化又影响到棋子的走法，如此循环下去。这里虽然有不同的层次之分，但是关于层次“高级”与“低级”的说法已经没有意义了。

不过我们再细细分析一下，仍然可以找出许多保持不变的东西。例如下棋要按一定的规则，对手们轮流下棋等等。也就是说存在着一种不变化的层次。因此我们可以把各种层次分为性质不同的两类。一类层次是不变的，我们称为I层次。另外一类层次是可以变化的，我们称为T层次。T层次是可以缠绕的，I层次则不行。I层次控制着T层次上发生的事件，T层次却无法影响I层次。

也许你会设想，可以用某一种方式使I层次与T层次缠绕起来。实现了这种“超级缠绕”就意味着I层次的变化。例如你和下棋的对手达成协议，把每人轮流下一次改为一方多下一次。但是这样的结果使得原来属于I层次的层次变成了T层次。与此同时必然会产生新的I层次凌驾于这种“超级缠绕”之上。这个过程会无休止地延伸下去。

我们可以设想这样一种情形。Z是作者T作品中的人物；同时T是作者E作品中的人物；E是作者

Z作品中的人物。这样的关系究竟能否实现呢？当然可以。不过得有一个花招。那就是这三个人都是另外一位作者H笔下的人物。这样作者H就是不变的层次，它在Z、T和E所在的层次之上，是不受他们的影响的。

埃舍尔的作品《画画的双手》（图22）就是这样的例子。这里的左手画出了右手，右手同时画出了左手。它们以这种绝妙的方式相互缠绕在一起。但是它们都是由埃舍尔的手画出来的。埃舍尔的手在这两只手的层次之上。

在我们的思维中，符号与符号相互作用，这种作用可以自相缠绕。而且这些符号还可以彼此引起内部的变化。有趣的是这种符号的自相缠绕依赖于在低层次上神经元的自相缠绕。不过这是两种性质完全不同的自相缠绕。前者是一种怪圈，后者却类似于反馈的回路。我们要细心留意这种差别。

饶有兴味的是在政府机构中同样有层次自相缠绕的现象。就拿法庭来讲，当两个人因为某个案例而争讼时，法庭可以作出裁决。这时法庭与争讼者分属不同的层次。但是如果法庭本身也卷入了法律事件，那就截然不同了。这时就形成了怪圈。当然也可以让更高级的法庭来进行裁决，摆脱这种怪圈。

但是如果连最高级的法庭也被卷入纠纷那又怎么样呢？美国的水门事件就是这样。议会要求总统服从最高法院的裁决，但是总统却可以加以否决。因为他有法律上的权力在某些情况下不服从最高法院的裁决。这样就产生了一个无法摆脱的怪圈。

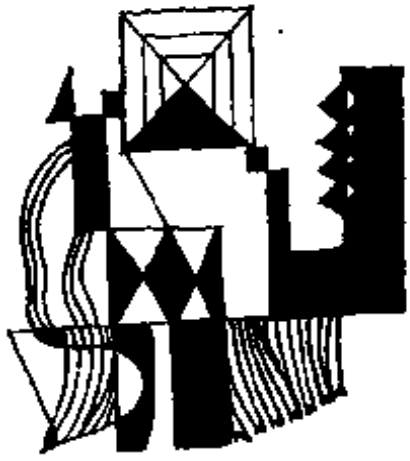
对于证明的评价也是一个复杂的问题。能否定义证明是什么呢？也许是做不到的。也许有人会想到希尔伯特对形式证明的定义。即严格按照形式系统的规则，从前提推到结论。但是我们已经看到，任何严格的规则都有例外，而不严格的规则又不能算是规则。其实不同的人、不同的人工智能程序都有不同的评价证明的标准。评价证明的问题实质上也就是判定真理性的问题，它在很大程度上是一种艺术，它依赖于对于美和简单性的意识就和依赖于逻辑原理和可以客观描述的依据一样。因此“什么是证明”并不是一个哲学问题，而是可以用不同层次加以解释的问题，它也渗透到各个领域里去。而当我们深入探讨这个问题时不可避免地会遇到怪圈。

当我们一开始见到埃舍尔画中的怪圈时，虽然感到有趣、迷惑，但是总有怪的感觉，感到这是一种不可思议的现象。现在我们可以看到，怪圈不仅不

是一种罕见的现象，而且在许多场合里是不可避免的。它是许多复杂系统的共同特点。既然如此，那么是否也存在“心理学的哥德尔定理”呢？如果我们着眼于领会这种说法的精神实质，而不是强求在形式上把这个定理翻译成心理学的术语，那么也许会给我们一些深刻的启示。我在这里仅仅提及两个例了。第一，把哥德尔定理和人的思维联系起来可以涉及到一个人是否明智的问题。你怎样判定自己是否明智呢?这是一个怪圈。你可以根据自己无矛盾的逻辑对世界作出某种解释，但是你怎么能判定自己的逻辑是否正确呢？当然只能根据自己的逻辑。我看不出怎样能找到问题的答案。哥德尔定理还有另一层意义上的隐喻。“我们无法理解自己的思维的。”这是一个需要谨慎分析的命题。怎样才叫“理解”呢？对此可以有种种不同的解释。对于“理解”的解释依赖于我们去理解的方法。也许禅宗的这种说法不无道理，我是无法理解我自己的。

9

结 束 语



这本书从怪圈开始回到了层次的自相缠绕。这本身就是一个巧妙编织而成的圈，是由各种怪圈构成的圈。在这个圈里我们讨论了形式系统、思维以及它们与现实世界的相互关系。我们的思维结构是以形式系统为基础的，是对现实世界的反映。但是禅宗的基本信条之一就是理性的思维是无法把握世界的，真理是无法用语言来描述的。

在禅宗的教义中包括大量悖论性质的内容。这些教义反对理性思维的逻辑而推崇超越理性的悟性。禅宗总是反对信赖语言的表达。在他们看来，

一旦用语言说出了某一个事物，就是在它和世界的其他部分之间划了一条界线，你就是人为地割裂了世界，于是就会陷入迷途。

可以说禅宗的教义是极端的总体论。总体论认为世界是结合成总体的，对于事物只能从总体上去把握，而不能简单地看成各部分之和。但是禅宗却走得更远。它们认为世界根本就不能分成各部分。一旦将世界分成各部分就会失去悟性。禅宗认为悟性就是这样一种状态，使醒悟者本身与世界其他部分之间的界线归于消失。可是对于一个人来讲，这种界限怎样才能消失呢？难道不就是死亡吗？

当然思维的逻辑并不是不可打破的。看了埃舍尔的杰作《白天和黑夜》（图15）就会有人要问："这是真正的飞鸟和田野吗？这是真实的白天和黑夜吗？"对于这些问题是无法按照思维的逻辑来回答的。埃舍尔很喜欢创造那些按照逻辑分析是自相矛盾的画。例如在《另一个世界》（图14）中，画面就以相同的形式来表现真实和不真实的世界。

《三个球》（图16）这幅画也是发人深省的。每个球作为世界的一部分都包含了世界的其他部分，同时又被包含在世界的其他部分中。我们可以在画面上看到每个球相互反映，也反映出书桌，

关于三个球的画以及画它们的画家。与此同时桌面又把它们统统反映出来。这暗示着世界上事物之间无穷的联系。这就象佛教故事中的“因陀罗之网”，它弥漫在整个宇宙中。它的水平线穿越空间，垂直线纵贯时间，交点则是每一个实体，它们是一颗颗晶粒。“神灵”的光辉照耀并穿透每颗晶粒。而每颗晶粒不但反射着来自每颗晶粒的光芒而且也反射着整个宇宙中互相反射的光芒。

于是在我的头脑中浮现了这样的图象：每一个重整化的电子中有虚的光子、质子、中子、介子……而在每个光子中又有虚的电子、质子、中子、介子……

我还想到了另一幅图景：每个人都反映其他许多人的思维，他们每个人又反映别人的思维，一个无穷无尽的系列……

对于自己会有一天不再存在的意识，也许是一种最大的矛盾。从一个层次上讲，当你脱离自己，把自己看成一个认识对象时，这种意识是完整的、很自然的。但是从另一个层次上讲，意识到的一切都是以某种方式嵌入我们的思维中来，否则就是不可理解的，因此个人的不存在就使这种意识变得毫无意义。

这种不相容性是两种观念的冲突。一种观念认为世界和我是统一的，关于我不再存在的说法只是术语上的矛盾。另一种观念则认为，我只是世界中的一部分，我是要死的，而世界没有我仍将继续存在下去。

科学在西方常常受到批评。说它二元论化。即主观与客观的分离，观察者与被观察对象的分离。但是现代物理学理论的发展正在走向两者的统一。而且随着科学的进展，人类正在越来越深入地研究人类思维，即观察机制本身。

和主客观二元论密切相关的是符号-对象的二元论。传统的美术和音乐都是用“符号”来表达思想和感情的。然而现代艺术却致力于发掘音乐和美术的这种能力，它们并不表示任何东西，而仅仅表现它们自身。这就意味着在艺术作品中可以有纯粹的色彩或者纯粹的声音，它们不起任何符号的作用，或者说是符号与对象融为一体。于是就有抽象派的绘画、“新古典主义”的音乐。

马格里特有一系列迷惑人心的语义画。我们以那幅《两件神秘的东西》（图21）为例。先看画面里的那幅画。你获得的信息是，这些符号与烟斗有区别。如果你把浮在空中的那个烟斗看成是真实

的，那么另一个就是一种符号。当然这一切是错误的。因为它们两者都是只是画面上的形象。如果把画中的画看成比画中另一部分更缺乏真实性这仅仅是一种视觉上的谬误。因为这样一来就会把画中的形象误认为是真实的东西。只有当你把这些都看成画面里的形象，你才能理解画上的那句题词："这不是烟斗"。可是在你接受这种观念的同时，这句话也成了带色的线条和斑块，从而失去了自己的意义。换言之，这幅画的语义信息是以哥德尔的方式自我瓦解的。

现实的世界是多层次的、无穷维的、不可分割的。既然如此，语言描述世界的能力有多大，思维反映世界的能力有多大呢？诚然，这种能力是一块永远可以无止境扩张的疆域。可是这也是一块永远有界限的疆域。无论是大脑思维，无论是人工智能，无论是严密的数学，无论是抽象的音乐，也无论是形象的美术，都是以形式系统作为自己最基础的层次。因而都无法摆脱奇妙的怪圈。

《音乐的奉献》表现了巴赫在旋律配合上最杰出的才华。它本身就是充满智慧的赋格，其中交织着各种思想和形式，它们的共同之处是双重含义和机敏的隐喻。这确实是人们永远可以欣赏的智慧杰

作。

也可以说没有人能真正看透《音乐的奉献》。当一个人自以为理解它的一切时突然又会有新的发现。例如在其中的六部无插入赋格的结尾，他不好意思地藏下了自己的名字。这里充满了音符和文字的游戏。这里有各种形式的卡农，有非常复杂的赋格、有美丽而深沉的情感，也有渗透各个层次的狂喜。它是赋格的赋格，是层次的自相缠绕。这和埃舍尔意味无穷的绘画相仿，与哥德尔深奥美妙的数学理论巧妙配合。这种结构使人想到无法用语言表达的人类思维的多音部赋格。这也就是为什么要在这本书中用哥德尔、埃舍尔和巴赫三股线编织出一条永恒的金带。